PRINCIPAL BUNDLES AND CURVATURE PART I; VECTOR BUNDLES

William Schulz
Department of Mathematics and Statistics
Northern Arizona University, Flagstaff, AZ 86011

1. INTRODUCTION

This[1] is a pdf with a treatment of Principal Bundles. My goal here is to make a treatment which is readily understandable by anyone with a year of graduate school in mathematics, but possibly this goal has not quite been met. At the moment this is α-ware, but I hope to improve and complete it over time.

There are many redundancies between sections so that it is easier to read without looking backward and forward.

The plan for this work is as follows. First we will present some elementary considerations which motivate our work and then present the theory of vector space bundles and curvature. Curvature will be presented in three ways; first with differential forms, then with Covariant Derivatives, and finally by using horizontal lifts of vector fields. We will show that all these methods result in the same curvature tensor. In the second part we will lift the results of the first part to principal bundles. We will first look at principal bundles as bundles of frames and then look at them in a more abstract setting. The third part will show how the theory for principal bundles can be used to give similar theories in all the associated bundles of the principal bundle. We will recover the original theory of part one as a special case. Finally in a fourth section we will apply the material to the case of Riemannian Geometry.

Although from time to time clever methods are used, in many cases the results are arrived at by direct calculation. These are presented in detail. Some effort is made to identify the important steps in the process. In some cases the calculations are not very much fun, and in these cases it is sometimes difficult to find them in other sources. At least I found it so. In most sources these places are identified by the words "An easy calculation shows......."

To read these notes you need to have some familiarity with exterior algebra and differential forms, differentiable manifolds and their tangent bundles. You also need to know how to represent vectors in the tangent bundle in terms of differential operators, so that i is represented by $\frac{\partial}{\partial x}$ etc. All this material can be found in other modules on this website which present material on Lie Derivatives and on Laplacians. Material currently unavailable here but important for these notes is the idea of the Lie Algebra of a Lie Group. This is commonly available material and we need little beyond the definition.

When reading these notes it would in many cases probably be counterproductive to read through all the calculations in detail; just concentrate on the

[1] November 17, 2015

ones you have some special interest in and glance over the others to get an idea of how they work.

2. HOW IT ALL STARTED–POLAR COORDINATES ON THE PLANE AND SPHERE

If we attempt to use polar coordinates in the plane and wish to take derivatives of vector fields written in terms of the polar coordinate base vectors complications will arise due to the fact that unlike the classical i, j vectors, the polar coordinate base vectors are *not* fixed. The changes in the base vectors must be built into the system somehow, and we show how to do this below. These fixes give rise to the notion of connection.

We use as usual $x = r\cos\theta$ and $y = r\sin\theta$ and $\{\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\}$ for our basis in polar coordinates and relate this to the fixed basis $\{i, j\} = \{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\}$. We will need the Jacobians

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

and

$$\frac{\partial(r,\theta)}{\partial(x,y)} = \frac{1}{r}\begin{pmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

We have

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta} \\ &= \cos\theta\,\frac{\partial}{\partial r} - \frac{1}{r}\cos\theta\,\frac{\partial}{\partial \theta} \\ \frac{\partial}{\partial y} &= \frac{\partial r}{\partial y}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial y}\frac{\partial}{\partial \theta} \\ &= \sin\theta\,\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\,\frac{\partial}{\partial \theta} \end{aligned}$$

which we can write more succinctly as

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) = \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right)\frac{\partial(r,\theta)}{\partial(x,y)} = \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right)\begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{r}\sin\theta & \frac{1}{r}\cos\theta \end{pmatrix}$$

$$\left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right) = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)\frac{\partial(x,y)}{\partial(r,\theta)} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

The vectors $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$ are essentially i, j, which are constant vectors. Suppose now $v = f^1\frac{\partial}{\partial r} + f^2\frac{\partial}{\partial \theta}$. To determine $\frac{\partial v}{\partial r}$ we rewrite in terms of $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$:

$$\begin{aligned} v &= \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right)\begin{pmatrix} f^1 \\ f^2 \end{pmatrix} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}\begin{pmatrix} f^1 \\ f^2 \end{pmatrix} \\ &= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)\begin{pmatrix} \cos\theta\, f^1 - r\sin\theta\, f^2 \\ \sin\theta\, f^1 + r\cos\theta\, f^2 \end{pmatrix} \end{aligned}$$

so

$$\begin{aligned}
\frac{\partial v}{\partial r} &= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)\begin{pmatrix} \cos\theta\,\frac{\partial f^1}{\partial r} - r\sin\theta\,\frac{\partial f^2}{\partial r} - \sin\theta\, f^2 \\ \sin\theta\,\frac{\partial f^1}{\partial r} + r\cos\theta\,\frac{\partial f^2}{\partial r} - \cos\theta\, f^2 \end{pmatrix} \\
&= \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right)\begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{r}\sin\theta & \frac{1}{r}\cos\theta \end{pmatrix} \times \\
&\quad \left[\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}\begin{pmatrix} \frac{\partial f^1}{\partial r} \\ \frac{\partial f^2}{\partial r} \end{pmatrix} + \begin{pmatrix} -\sin\theta\, f^2 \\ \cos\theta\, f^2 \end{pmatrix}\right] \\
&= \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right)\left[\begin{pmatrix} \frac{\partial f^1}{\partial r} \\ \frac{\partial f^2}{\partial r} \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{r} f^2 \end{pmatrix}\right]
\end{aligned}$$

The first term is exactly what we would expect but the second term is a surprise to the uninitiated. If we look back through the equations we see that it's origin is due to the fact that the base vectors $\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}$ are *moving,* and thus a correction term must be added to compensate for this movement. The coefficients in the correction terms are called Christoffel symbols.

We can perform the above in a more systematic way. To minimize notational horror we will use e_1, e_2 for $\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}$ and i, j for $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$. Then

$$\begin{aligned}
(e_1, e_2) &= (i, j)\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \\
\frac{\partial}{\partial r}(e_1, e_2) &= (i, j)\begin{pmatrix} 0 & -\sin\theta \\ 0 & \cos\theta \end{pmatrix} \\
&= (e_1, e_2)\begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{r}\sin\theta & \frac{1}{r}\cos\theta \end{pmatrix}\begin{pmatrix} 0 & -\sin\theta \\ 0 & \cos\theta \end{pmatrix} \\
&= (e_1, e_2)\begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{r} \end{pmatrix}
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial}{\partial \theta}(e_1, e_2) &= (i, j)\begin{pmatrix} -\sin\theta & -r\cos\theta \\ \cos\theta & -r\cos\theta \end{pmatrix} \\
&= (e_1, e_2)\begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{r}\sin\theta & \frac{1}{r}\cos\theta \end{pmatrix}\begin{pmatrix} -\sin\theta & -r\cos\theta \\ \cos\theta & -r\sin\theta \end{pmatrix} \\
&= (e_1, e_2)\begin{pmatrix} 0 & -r \\ \frac{1}{r} & 0 \end{pmatrix}
\end{aligned}$$

From this

$$\begin{aligned}
\frac{\partial}{\partial r} v &= \frac{\partial}{\partial r}(e_i f^i) = e_i\,\frac{\partial f^i}{\partial r} + \frac{\partial e_i}{\partial r} f^i \\
&= (e_1, e_2)\begin{pmatrix} \frac{\partial f^1}{\partial r} \\ \frac{\partial f^2}{\partial r} \end{pmatrix} + (e_1, e_2)\begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{r} \end{pmatrix}\begin{pmatrix} f^1 \\ f^2 \end{pmatrix} \\
&= (e_1, e_2)\begin{pmatrix} \frac{\partial f^1}{\partial r} \\ \frac{\partial f^2}{\partial r} \end{pmatrix} + e_2\,\frac{1}{r} f^2
\end{aligned}$$

as before. Now however $\partial v/\partial\theta$ is much simpler to get:

$$\begin{aligned}
\frac{\partial}{\partial\theta}\, v &= (e_1, e_2)\begin{pmatrix} \frac{\partial f^1}{\partial\theta} \\ \frac{\partial f^2}{\partial\theta} \end{pmatrix} + (e_1, e_2)\begin{pmatrix} 0 & -r \\ \frac{1}{r} & 0 \end{pmatrix}\begin{pmatrix} f^1 \\ f^2 \end{pmatrix} \\
&= (e_1, e_2)\begin{pmatrix} \frac{\partial f^1}{\partial\theta} \\ \frac{\partial f^2}{\partial\theta} \end{pmatrix} + (e_1, e_2)\begin{pmatrix} -rf^2 \\ \frac{1}{r} f^1 \end{pmatrix}
\end{aligned}$$

This gives the flavor of the situation but there is one more complication which we will illustrate by doing similar things to the example of the sphere embedded in three space. What makes the polar coordinate example over-special is that the derivatives of the base vectors $\partial/\partial r, \partial/\partial\theta$ are linear combinations of these same vectors. This will not be true for the sphere, and we will have to modify our concept of derivative to compensate for this circumstance. Put another way, the ordinary derivatives of vectors in the tangent bundle of the sphere embedded in three space are not in the tangent bundle.

We can parametrize the sphere by the angle ϕ measured off the z-axis (often called colatitude) and the angle θ of the vertical plane through the point and the origin off the x-axis called longitude.

The position vector (in three space) of a point on the sphere is given by

$$\mathbf{x}(\phi, \theta) = < a\sin\phi\cos\theta, a\sin\phi\sin\theta, a\cos\phi >$$

Then the base vectors for the tangent bundle and the unit normal are

$$\begin{aligned}
e_1 &= \frac{\partial \mathbf{x}}{\partial\phi} = < a\cos\phi\cos\theta, a\cos\phi\sin\theta, -a\sin\phi > \\
e_2 &= \frac{\partial \mathbf{x}}{\partial\theta} = < -a\sin\phi\sin\theta, a\sin\phi\cos\theta, 0 > \\
n_1 &= e_1 \times e_2 = < a^2\sin^2\phi\cos\theta, a^2\sin^2\phi\sin\theta, a^2\sin\phi\cos\theta > \\
&= a^2\sin\phi\ < \sin\phi\cos\theta, \sin\phi\sin\theta, \cos\theta > \\
n &= < \sin\phi\cos\theta, \sin\phi\sin\theta, \cos\theta >
\end{aligned}$$

where n is the unit tangent vector. Then, for some coefficients Γ^k_{ij}

$$\frac{\partial e_i}{\partial u^j} = e_k \Gamma^k_{ij} + b_{ij}\, n$$

Notice that the term $b_{ij}\, n$ carries us outside the tangent bundle. We do not want to consider things happening outside this bundle, so we project $\partial e_i/\partial u_j$ back into the bundle. We will denote this modified derivative by

$$\frac{De_i}{\partial u^j} = e_k \Gamma^k_{ij}$$

or, written another way,

$$De_i = e_k \Gamma^k_{ij}\, du^j = e_k\, {\omega_i}^k$$

The matrix ${\omega_i}^k$ is called the *connection matrix* and the coefficients Γ^k_{ij} are called the *Christoffel symbols.* The operator D is called the *covariant differential.*

For the curious we will present the actual coefficients in this case; the user can verify that they are correct.

$$\begin{aligned}
\frac{D}{\partial\phi}(e_1,e_2) &= (e_1,e_2)\begin{pmatrix} \Gamma^1_{11} & \Gamma^1_{21} \\ \Gamma^2_{11} & \Gamma^2_{21} \end{pmatrix} = (e_1,e_2)\begin{pmatrix} 0 & 0 \\ 0 & \cot\phi \end{pmatrix} \\
\frac{D}{\partial\theta}(e_1,e_2) &= (e_1,e_2)\begin{pmatrix} \Gamma^1_{12} & \Gamma^1_{22} \\ \Gamma^2_{12} & \Gamma^2_{22} \end{pmatrix} = (e_1,e_2)\begin{pmatrix} 0 & -\sin\phi\cos\phi \\ \cot\phi & 0 \end{pmatrix}
\end{aligned}$$

To check these it is handy to have

$$(b_{ij}) = \begin{pmatrix} -a & 0 \\ 0 & -a\sin^2\phi \end{pmatrix}$$

so, for example,

$$\begin{aligned}
\frac{\partial e_2}{\partial u^2} &= e_k\Gamma^k_{22} + b_{22}\,n \\
&= e_1(-\sin\phi\cos\phi) - a\sin^2\phi\,n
\end{aligned}$$

In the general case of manifolds which are not embedded it is necessary to supply the connection forms or Christoffel symbols from outside; there is no natural choice unless the manifold has additional structures, for example a Riemannian structure. Lacking this, we can supply a connection by choosing a local basis of sections and coefficients Γ^k_{ij} *with respect to this basis* using some external criterion, for example additional geometry or some physical situation or pure whimsy. There is also an abstract definition which covers all possible choices. Once the coefficients have been chosen for some choice of local basis of sections they are determined for all others. Also, while we have worked within the tangent bundle, the same methods will work on any other vector bundle.

3. FIBRE BUNDLES

We want to quickly define Fibre Bundles. No effort has been made to develop the general theory; we are only interested in Vector Bundles and Principal Bundles. But the general case puts it all in a sort of context.

Let M be a differentiable manifold and $\pi : \mathcal{E} \to M$ be a submersion (π has maximal rank). For "small" subsets U of M we want $\pi^{-1}[U]$ to be isomorphic to a product of U and a fixed topological space E, called the *fibre,* and this part of $\mathcal{E}$ we think of as "locally trivial" since it is just a product. Globally $\mathcal{E}$ is *not* such a product, of course. To do this we assume that for each $x \in M$ there exists a neighborhood U of x so that $\pi^{-1}[U]$ is homeomorphic to a product (with fixed E). That is, there exists a map ϕ called a local trivialization for which

$$\phi : \pi^{-1}[U] \to U \times E$$

and a decomposition of π as $\pi(p) = \pi_1 \circ \phi(p)$, where $p \in \mathcal{E}$ and π_1 is projection on the first component of the product.

If x happens to be in two of the trivializing neighborhoods U and V then $\phi_{VU} : E \to E$ is defined, for $e \in E$, by

$$(x, \phi_{VU}(e)\,) = \phi_V \circ \phi_U^{-1}(x, e)$$

It is assumed that ϕ_{VU} is a homeomorphism, and also a diffeomorphism if E is a differentiable manifold or an isomorphism if E is a vector space or an automorphism if E is a group. In general ϕ_{VU} will be an equivalence of E in the relevant category.

The ϕ_{VU} are called *transition functions*. They must satisfy the equation, for $x \in U \cap V \cap W$

$$\phi_{WV} \circ \phi_{VU} = \phi_{WU}$$

There is a more elegant way to look at this: we first note that $\phi_{VU} = \phi_{UV}^{-1}$. Then we can write the above equation as

$$\phi_{WV} \circ \phi_{VU} \circ \phi_{UW} = id_{U \cap V \cap W} \qquad \text{The cocycle condition}$$

Although we will not make use of it, if you know the fibre E, the cover $\{U, V, W, \ldots\}$ and the transition functions ϕ_{VU} you can reconstruct the fibre bundle by standard gluing techniques from topology. The name *cocycle condition* comes from Čech Cohomology.

The set of equivalences of E form a Lie group, which need not be a finite dimensional, but in the situations in which we will work the ϕ_{VU} will usually belong to a subgroup of $\mathrm{GL}(n, \mathbb{R})$ (or $\mathrm{GL}(n, \mathbb{C})$, (the symbolism is identical) and thus will indeed be a finite dimensional Lie Group. The group to which all the transition functions belong is called the *Structure Group*. If the bundle has additional structure, for example a vector bundle in which the fibres have a continuously varying inner product, it may be possible to find transition functions which belong to a proper subgroup of $\mathrm{GL}(n, \mathbb{R})$, for example $\mathrm{O}(n, \mathbb{R})$. This is called *reducing the structure group*. Popular choices of structure group are $\mathrm{GL}(n, \mathbb{R})$, $\mathrm{O}(n, \mathbb{R})$ which is the Orthogonal Group, or $\mathrm{SO}(n, \mathbb{R})$, which is the Special Orthogonal Group.

A *local section* over U is a smooth map $\rho : U \to \mathcal{E}$ such that $\pi \circ \rho : U \to U$ is the identity. Local sections are very common and important.

The complete notation for a Fibre Bundle is

$$(\mathcal{E}, M, \pi, E, G)$$

where $\mathcal{E}$ is the *total space*, M is the *base space*, π is the projection $\pi : \mathcal{E} \to M$, E is the *fibre*, and G is the *structure group*. We will often adhere to the convention that the total space is denoted by the calligraphic form of the fibre, but will not use this convention for principal bundles.

4. VECTOR BUNDLES

A Vector Bundle is a special case of a fibre bundle where the fibre E is a vector space and the transition functions ϕ_{VU} are vector space isomorphisms. (Note that this requirement on ϕ_{VU} is built into the definition of a vector bundle.) In this case there is a special way of viewing the trivializations which is very useful. Consider a trivialization $\phi : \pi^{-1}[U] \to U \times E$ where we can regard E as $\mathbb{R}^n$ (or $\mathbb{C}^n$; nothing changes in the formalism). Then $\sigma_\alpha = \phi^{-1}(x, (0, \ldots, 0, 1, 0, \ldots, 0))$, with 1 in the α^{th} place, is a section of $\mathcal{E}$ over U. Moreover, since $\phi_x : \pi^{-1}[x] \to E$ is an isomorphism, the set of sections $\sigma_1(x), \ldots, \sigma_n(x)$ forms a basis for the fibre $\pi^{-1}[x]$, and thus $\sigma_1, \ldots, \sigma_n$ is a local basis of sections for $\mathcal{E}$ over U. Any differentiable section $\rho : U \to \mathcal{E}$ can then be written as

$$\rho = \sigma_\alpha \rho^\alpha$$

where the ρ^α are differentiable functions on the coordinate patch U, and moreover

$$\phi(\rho(x)) = (x, \rho^1(x), \ldots, \rho^n(x))$$

Clearly the procedure may be reversed; given a local basis σ_α of smooth sections (that is, $\sigma_1(x), \ldots, \sigma_n(x)$ are linearly independent for each $x \in U$) we can construct a corresponding trivialization over U by

$$\phi(p) = (x, \rho^1(x), \ldots, \rho^n(x))$$

where $\pi(p) = x$ and $\rho(x) = \sigma_\alpha(x)\rho^\alpha(x)$. Using the above procedure on the trivialization gives us back the same local basis of sections σ_α that we started with.

Naturally when we work with vector bundles the structure group in which the ϕ_{VU} live will be $\mathrm{GL}(n, \mathbb{R})$ or some subgroup of it, as discussed at the end of the previous section.

5. PRINCIPAL BUNDLES

This section logically goes at this place, but the material in it will not be used until Part II of the notes. Hence the reader may wish to put off reading this section until she gets to Part II.

For Principal Bundles there are some slight modifications of the notation. What characterizes Principal bundles is that the fibre E is homeomorphic to the structure group G which is required to be a topological group (and for this work a Lie Group). Thus the notation will be (P, M, π, G) where P is used instead of $\mathcal{E}$ for the total space and the fibre is omitted.

In addition, we require there to be right action of G on each fibre $\pi^{-1}[x]$, $x \in M$, and this right action must be free and transitive. Free means that if $p \in \pi^{-1}[x]$ and $pg_1 = pg_2$ then $g_1 = g_2$. Transitive means that if $p_1, p_2 \in \pi^{-1}[x]$ then there exists a $g \in G$ for which $p_2 = p_1 g$. This g is unique by freeness. Thus for any $p \in \pi^{-1}[x]$, the entire fibre $\pi^{-1}[x]$ is precisely the orbit of p under the G action, and moreover by freeness we have a one to one correlation

$$\begin{array}{rcl} \pi^{-1}[x] & \longleftrightarrow & G \\ pg & \longleftrightarrow & g \end{array}$$

However, and this is most important, this correlation is not canonical; it depends on the choice of $p \in \pi^{-1}[x]$. However, one often speaks loosely and says that the fibre is the structure group G. (Technically, the fibres are G-torsors. A G-torsor is a topological space homeomorphic with the topological group G but not necessarily having any group structure or identity. Of course we can always *put* these on the torsor, but there is usually no canonical way to do it.)

As with any fibre bundle we have transition functions. Let the standard homeomorphism be

$$\begin{array}{ll} \Phi_U : \pi^{-1}[U] \to U \times G & \Phi_U(p) = (\pi(p), g_U(p)) \\ \Phi_V : \pi^{-1}[V] \to V \times G & \Phi_V(p) = (\pi(p), g_V(p)) \end{array}$$

and then we have the transition function on $U \cap V$

$$\begin{array}{rcl} \Phi_{VU} & = & \Phi_V \circ \Phi_U^{-1} \\ \Phi_{VU}(\pi(p), g_U(p)) & = & (\pi(p), g_V(p)) \end{array}$$

Here g_U is a function from $\pi^{-1}[U]$ to G and for fixed $x \in M$ it determines a homeomorphism between $\pi^{-1}[x]$ and G. There is a similar homeomorphism determined by g_V. There is also a function g_{VU} from $\pi^{-1}[U \cap V]$ to G corresponding to Φ_{VU} given by $g_{VU}(p) = g_V(p) g_U(p)^{-1}$. We thus have

$$\Phi_{VU}(\pi(p), g_U(p)) = (\pi(p), g_{VU}(p) g_U(p))$$

If W is a third coordinate patch and $U \cap V \cap W \neq \emptyset$ Then the transition functions must satisfy

$$g_{WV} \circ g_{VU} = g_{WU} \quad \text{on} \quad U \cap V \cap W$$

which, using $g_{UW} = g_{WU}^{-1}$ we can write more elegantly as

$$g_{WV} \circ g_{VU} \circ g_{UW} = Id \quad \text{on} \quad U \cap V \cap W$$

Finally we note that given a vector bundle $(\mathcal{E}, M, \pi, E, G)$ we will have a set of coordinate patches and transition functions Φ_{UV}. Using M, the coordinate patches, the transition functions and the group G we can construct a principal bundle (P, M, π, G) by standard gluing techniques of topology. The original bundle is then said to be an *associated bundle* of the principal bundle, and we will eventually learn how to construct a multitude of associated G-bundles to P. We will go into this in excruciating detail in later sections.

This position of P at the center of a family of associated bundles means that P functions as a sort of Überbundle. For us, the important point is that is that once we have lifted a connection from a vector bundle to its principal bundle, the connection metastasizes automatically to each of the associated bundles.

Canonical Trivializations Functions of a Principal Bundle

For some purposes it is useful to have local trivialization functions Φ_U for a principal bundle which satisfy certain additional conditions. We will call these *canonical trivializations* although this terminology is not standard. We will use $\tilde{\Phi}_U$ for these special trivializations.

A canonical trivialization has the property that, with $h \in G$,

$$\begin{array}{lrcl} \text{If} & \tilde{\Phi}_U(p) & = & (\pi(p), g_U(p)) \\ \text{Then} & \tilde{\Phi}_U(ph) & = & (\pi(p), g_U(p)h) \end{array}$$

which can be written as $g_U(ph) = g_U(p)h$. This says that the trivialization functions respect the action of G on P.

We may construct canonical trivialization functions from ordinary trivialization functions easily as follows. Let U be a coordinate patch and $\Phi_U : \pi^{-1}[U] \to G$ be an ordinary trivialization. Let $\sigma_U : U \to P$ be a local section over U. Set

$$\tilde{\Phi}_U(\sigma_U(x)) = \Phi_U(\sigma_U(x)) = (x, g_x) \quad \text{for} \quad x \in U$$

Then for any $p \in \pi^{-1}[U]$ there exists a unique $h_p \in G$ for which $p = \sigma_U(x)h$, by transitivity and freeness. Now define

$$\tilde{\Phi}_U(p) = (x, g_x h_p) = (x, g_p)$$

with $g_p = g_x h_p$. I claim that $\tilde{\Phi}_U$ is a canonical trivialization. Indeed, $ph = \sigma(x)h_p h$ so

$$\tilde{\Phi}_U(ph) = (x, g_x h_p h) = (x, g_p h)$$

as required.

Next we discuss the transition functions associated with canonical trivializations and the associated functions g_U, g_V, g_{VU}. This is easy but a bit clumsy. We will be using the notation of the first part of this section.

Let $(x, g) \in U \times G$. For any x and g there will exist a $p \in P$ for which

$$(x, g) = \tilde{\Phi}_U(p) = (\pi(p), g_U(p))$$

Since $\tilde{\Phi}_U$ is canonical we have

$$(x, gh) = \tilde{\Phi}_U(ph) = (\pi(p), g_U(p)h)$$

but since in general $\tilde{\Phi}_U(ph) = (\pi(p), g_U(ph))$ we have

$$g_U(ph) = g_U(p)h$$

Next we see that

$$\tilde{\Phi}_U^{-1}(x, g) = p \qquad \tilde{\Phi}_U^{-1}(x, gh) = ph$$

and then

$$\tilde{\Phi}_{VU}(x,g) = \tilde{\Phi}_V(\tilde{\Phi}_U^{-1}(x,g)) = \tilde{\Phi}_V(p) = (x, g_V(p))$$

$$\tilde{\Phi}_{VU}(x,gh) = \tilde{\Phi}_V(\tilde{\Phi}_U^{-1}(x,gh)) = \tilde{\Phi}_V(ph) = (x, g_V(p)h)$$

because $\tilde{\Phi}_V$ is canonical. This shows that $\tilde{\Phi}_{VU}$ respects the right action of G on P.

Now what happens with g_{VU}? We have

$$\begin{aligned} g_{VU}(ph) &= g_V(ph)g_U(ph)^{-1} \\ &= g_V(p)h[g_U(p)h]^{-1} \\ &= g_V(p)hh^{-1}g_U(p) \\ &= g_V(p)g_U(p) \\ &= g_{VU}(p) \end{aligned}$$

This tells us that the functions $g_{VU}(p)$ are constant on each fibre $\pi^{-1}[x]$ so that they may be written as $g_{VU}(x)$. Then we notice that, with $(x,g) = \tilde{\Phi}_U(p)$,

$$\tilde{\Phi}_{VU}(x,g) = (x, g_V(p)) = (x, g_{VU}(x)g_U(p)) = (x, g_{VU}(x)g)$$

so the transition functions are essentially left actions of G on each fibre. If we combine this with the right action we get

$$\tilde{\Phi}_{VU}(x,gh) = (x, g_V(ph)) = (x, g_{VU}(x)g_U(p)h) = (x, g_{VU}(x)gh)$$

so that we have commuting left and right actions of G, which commute due to the associative law on G.

6. CONNECTION AND CURVATURE FORMS ON VECTOR BUNDLES

In this section we introduce the connection and curvature forms and set our standards for notation. This section is for quick reference. Motivation and applications are handled at length in later sections.

We are dealing here with an n-dimensional Bundle $\mathcal{E}$ over a d-dimensional manifold M. Indices $i, j, k, \cdots$ run from 1 to d and indices $\alpha, \beta, \gamma, \cdots$ run from 1 to n

A connection is given by a set of 1-forms. As such they can be expressed (over U) in terms of the local coordinates du^i. The coefficients can be quite arbitrary. Thus a connection form is given by

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

The connection form gives rise to a curvature 2-form on the manifold. We will show in subsequent sections that the curvature form is

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

$$
\begin{aligned}
&= d(\omega_\beta{}^\alpha) + (\omega_\gamma{}^\alpha) \wedge (\omega_\beta{}^\gamma) \\
&= d(\Gamma^\alpha_{\beta l}\, du^l) + (\Gamma^\alpha_{\gamma k}\, du^k) \wedge (\Gamma^\gamma_{\beta l}\, du^l) \\
&= \Big(\frac{\partial \Gamma^\alpha_{\beta l}}{\partial u^k} + \Gamma^\alpha_{\gamma k}\, \Gamma^\gamma_{\beta l}\Big)\, du^k \wedge du^l
\end{aligned}
$$

If we antisymmetrize this we get

$$
\begin{aligned}
\Omega &= (\Omega_\beta{}^\alpha) \\
&= \frac{1}{2}\Big(\frac{\partial \Gamma^\alpha_{\beta l}}{\partial u^k} - \frac{\partial \Gamma^\alpha_{\beta k}}{\partial u^l} + \Gamma^\alpha_{\gamma k}\, \Gamma^\gamma_{\beta l} - \Gamma^\alpha_{\gamma l}\, \Gamma^\gamma_{\beta k}\Big)\, du^k \wedge du^l \\
&= \frac{1}{2}\, R_\beta{}^\alpha{}_{kl}\, du^k \wedge du^l \\
&= \sum_{k<l} R_\beta{}^\alpha{}_{kl}\, du^k \wedge du^l
\end{aligned}
$$

where we set

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

7. WHERE CONNECTIONS COME FROM

Connections are a way of differentiating sections of a vector bundle. They are also a way of identifying the fibres at different points along curves. There are many ways to introduce them and there appears no optimal choice. Our method is commonly used.

We define an operator D:

$$
D : \Gamma(\mathcal{E}) \to \Gamma(\mathcal{E}) \otimes \Gamma(T(M)^*)
$$

which is required to satisfy four properties. The first two are linearity and Leibniz' rule for sections of $\Gamma(\mathcal{E})$

$$
\begin{aligned}
&D(\rho + \tau) = D\rho + D\tau \\
&D(f\rho) = df \otimes \rho + f D\rho
\end{aligned}
$$

The second two are the tensorial properties for inputs from $\Gamma(T(M))$:

$$
\begin{aligned}
&(D\rho)(X + Y) = (D\rho)(X) + (D\rho)(Y) \\
&(D\rho)(fX) = f(D\rho)(X)
\end{aligned}
$$

In the future we will write $(D\rho)(X)$ as $D\rho(X)$. Emotionally one thinks that for a section $\rho \in \Gamma(\mathcal{E})$, $D\rho$ is an operator taking vectors $X \in T(M)$ as input and outputting sections, which are supposed to be the directional derivative of ρ in the direction X.

There is another way of viewing these laws which is also very useful. We first define

$$
D_X \rho = D\rho(X)
$$

and then, using $df(X) = Xf$, we can rewrite the above four laws in the form

$$\begin{aligned}
&D_X(\rho+\tau) = D_X(\rho) + D_X(\tau)\\
&D_X(f\rho) = X(f)\rho + fD_X(\rho)\\
&D_{X+Y}(\rho) = D_X(\rho) + D_Y(\rho)\\
&D_{fX}(\rho) = fD_X(\rho)
\end{aligned}$$

Only the second requires a little explanation:

$$\begin{aligned}
D_X(f\rho) &= D(f\rho)(X) = (df\otimes\rho + fD\rho)(X)\\
&= df(X)\rho + fD(\rho)(X)\\
&= X(f)\rho + fD_X(\rho)
\end{aligned}$$

8. FORMULAS IN A COORDINATE PATCH AND THE COVARIANT DERIVATIVE

If we want to actually create a connection or work with one in a coordinate patch, here is one method for doing it. We have coordinates $u^1, \cdots, u^d$ in a coordinate patch $U \subseteq M$ and a local basis of sections $e_1, \cdots, e_d$ for $\Gamma(T(M))$ and a basis $\sigma_1, \cdots, \sigma_n$ of local sections of the bundle $\mathcal{E}$. Then to get the connection locally we need only specify

$$D_{e_i}(\sigma_\alpha) = \sigma_\beta \Gamma^\beta_{\alpha i}$$

which should be thought of as the directional derivative of σ_α in the direction specified by e_i. Once we have the $\Gamma^\beta_{\alpha i}$ we can get all the other connection stuff by tensor laws and Leibniz' rule as I now demonstrate. It would probably be useful to keep in mind that one choice for $e_1 \cdots, e_d$ would be the natural basis $\frac{\partial}{\partial u^1} \cdots, \frac{\partial}{\partial u^d}$.

Let $X = e_i X^i$ and $\rho = \sigma_\alpha \rho^\alpha$ be the local coordinate representations of $X \in \Gamma(T(M))$ and $\rho \in \Gamma(E)$. Then

$$\begin{aligned}
D_{e_i}(\rho) &= D_{e_i}(\sigma_\alpha\rho^\alpha)\\
&= \sigma_\alpha D_{e_i}(\rho^\alpha) + D_{e_i}(\sigma_\alpha)\rho^\alpha\\
&= \sigma_\alpha e_i(\rho^\alpha) + \sigma_\beta \Gamma^\beta_{\alpha i}\rho^\alpha\\
&= \sigma_\beta\big(e_i(\rho^\beta) + \Gamma^\beta_{\alpha i}\rho^\alpha\big)\\
D_X\rho &= \sigma_\beta e_i(\rho^\beta)X^i + \sigma_\beta\Gamma^\beta_{\alpha i}\rho^\alpha X^i\\
&= \sigma_\beta\big(X(\rho^\beta) + \Gamma^\beta_{\alpha i}\rho^\alpha X^i\big)
\end{aligned}$$

For calculational purposes one often takes $X = \frac{\partial}{\partial u^j} = \frac{\partial}{\partial u^i}\delta^i_j$ and the formula then becomes

$$\begin{aligned}
D_{\frac{\partial}{\partial u^j}}\rho &= \sigma_\beta\Big(\frac{\partial\rho^\beta}{\partial u^j} + \Gamma^\beta_{\alpha i}\rho^\alpha\delta^i_j\Big)\\
&= \sigma_\beta\Big(\frac{\partial\rho^\beta}{\partial u^j} + \Gamma^\beta_{\alpha j}\rho^\alpha\Big)
\end{aligned}$$

Because the combination in parentheses in the last line occurs so frequently there is an ancient notation for it;

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

Often the notation is $\rho^\beta{}_{;j}$ instead of $\rho^\beta{}_{|j}$. We can then rewrite the previous equation as

$$D_{\frac{\partial}{\partial u^j}}\rho = \sigma_\beta\ \rho^\beta{}_{|j}$$

and if $X = X^j\frac{\partial}{\partial u^j}$ then

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

D_X is called a *covariant derivative.* Sometimes $D_{\frac{\partial}{\partial u^j}}$ is abbreviated D_j.

9. MIXED COVARIANT DERIVATIVE FORMULA

The covariant derivative D_X has many properties analogous to a partial derivative, but unlike ordinary partial derivatives, covariant derivatives do not commute. This is very important since the Curvature Tensor measures the deviation from commuting. Put another way, the failure of covariant derivatives to commute and the curvature of space are two sides of the same coin.

Let $e_1 = \frac{\partial}{\partial u^1}, \cdots, e_d = \frac{\partial}{\partial u^d}$ be a local basis for the tangent space and $\sigma_1, \cdots, \sigma_n$ be a local basis for the bundle.

Recall

$$D_{e_j}\sigma_\alpha = \sigma_\gamma\Gamma^\gamma_{\alpha j}$$

so for a tangent vector $Y = e_jY^j$ we have

$$D_Y\sigma_\alpha = D_{e_j}\sigma_\alpha Y^j = \sigma_\gamma\Gamma^\gamma_{\alpha j}Y^j$$

and finally, for a local section $\rho = \sigma_\alpha\rho^\alpha$ of the bundle we have

$$\begin{aligned} D_Y\rho &= \Big[D_{e_j}\big(\sigma_\alpha\rho^\alpha\big)\Big]Y^j \\ &= \Big[\sigma_\gamma\big(\frac{\partial\rho^\gamma}{\partial u^j} + \Gamma^\gamma_{\alpha j}\rho^\alpha\big)\Big]Y^j \\ &= \sigma_\gamma\rho^\gamma{}_{|j}\ Y^j \\ &= \sigma_\gamma M^\gamma \end{aligned}$$

where

$$M^\gamma = \rho^\gamma{}_{|j}\ Y^j$$

We are now going to deal with repeated covariant derivatives. This material is technically annoying but full details are provided here and things are treated extremely systematically (and thus not quite maximally efficiently).

$$\begin{aligned}
D_X D_Y \rho &= D_X(\sigma_\beta M^\beta) = \sigma_\gamma \Big(\frac{\partial M^\gamma}{\partial u^i} + \Gamma^\gamma_{\alpha i} M^\alpha \Big) X^i \\
&= \sigma_\gamma \left(\frac{\partial}{\partial u^i} \Big(\Big[\frac{\partial \rho^\gamma}{\partial u^j} + \Gamma^\gamma_{\beta j} \rho^\beta \Big] Y^j \Big) + \Gamma^\gamma_{\alpha i} \Big(\frac{\partial \rho^\alpha}{\partial u^j} + \Gamma^\alpha_{\beta j} \rho^\beta \Big) Y^j \right) X^i \\
&= \sigma_\gamma \Bigg(\frac{\partial^d \rho^\gamma}{\partial u^i \partial u^j} Y^j + \frac{\partial \Gamma^\gamma_{\beta j}}{\partial u^i} \rho^\beta Y^j + \Gamma^\gamma_{\beta j} \frac{\partial \rho^\beta}{\partial u^i} Y^j + {\rho^\gamma}_{|j} \frac{\partial Y^j}{\partial u^i} \\
&\quad + \Gamma^\gamma_{\alpha i} \frac{\partial \rho^\alpha}{\partial u^j} Y^j + \Gamma^\gamma_{\alpha i} \Gamma^\alpha_{\beta j} Y^j \rho^\beta \Bigg) X^i \\
&= \sigma_\gamma \left(\frac{\partial^d \rho^\gamma}{\partial u^i \partial u^J} + \frac{\partial \Gamma^\gamma_{\beta j}}{\partial u^i} \rho^\beta + \Gamma^\gamma_{\alpha i} \Gamma^\alpha_{\beta j} \rho^\beta \right) X^i Y^j \\
&\quad + \sigma_\gamma \left(\Gamma^\gamma_{\beta j} \frac{\partial \rho^\beta}{\partial u^i} X^i Y^j + \Gamma^\gamma_{\alpha i} \frac{\partial \rho^\alpha}{\partial u^j} X^i Y^j \right) + \sigma_\gamma \left({\rho^\gamma}_{|j} \frac{\partial Y^j}{\partial u^i} X^i \right) \\
D_Y D_X \rho &= \sigma_\gamma \left(\frac{\partial^d \rho^\gamma}{\partial u^i \partial u^J} + \frac{\partial \Gamma^\gamma_{\beta j}}{\partial u^i} \rho^\beta + \Gamma^\gamma_{\alpha i} \Gamma^\alpha_{\beta j} \rho^\beta \right) Y^i X^j \\
&\quad + \sigma_\gamma \left(\Gamma^\gamma_{\beta j} \frac{\partial \rho^\beta}{\partial u^i} Y^i X^j + \Gamma^\gamma_{\alpha i} \frac{\partial \rho^\alpha}{\partial u^j} Y^i X^j \right) + \sigma_\gamma \left({\rho^\gamma}_{|j} \frac{\partial X^j}{\partial u^i} Y^i \right) \\
&= \sigma_\gamma \left(\frac{\partial^d \rho^\gamma}{\partial u^j \partial u^i} + \frac{\partial \Gamma^\gamma_{\beta i}}{\partial u^j} \rho^\beta + \Gamma^\gamma_{\alpha j} \Gamma^\alpha_{\beta i} \rho^\beta \right) Y^j X^i \\
&\quad + \sigma_\gamma \left(\Gamma^\gamma_{\beta i} \frac{\partial \rho^\beta}{\partial u^j} Y^j X^i + \Gamma^\gamma_{\alpha j} \frac{\partial \rho^\alpha}{\partial u^i} Y^j X^i \right) + \sigma_\gamma \left({\rho^\gamma}_{|i} \frac{\partial X^i}{\partial u^j} Y^j \right) \\
\big(D_X D_Y - D_Y D_X \big) \rho &= \sigma_\gamma \left(\frac{\partial \Gamma^\gamma_{\beta j}}{\partial u^i} - \frac{\partial \Gamma^\gamma_{\beta i}}{\partial u^j} + \Gamma^\gamma_{\alpha i} \Gamma^\alpha_{\beta j} - \Gamma^\gamma_{\alpha j} \Gamma^\alpha_{\beta i} \right) \rho^\beta X^i Y^j \\
&\quad + \sigma_\gamma \left({\rho^\gamma}_{|j} \frac{\partial Y^j}{\partial u^i} X^i - {\rho^\gamma}_{|i} \frac{\partial X^i}{\partial u^j} Y^j \right) \\
&= \sigma_\gamma {R_\beta}^\gamma{}_{ij} \rho^\beta X^i Y^j + \sigma_\gamma {\rho^\gamma}_{|i} \left(X^j \frac{\partial Y^i}{\partial u^j} - Y^j \frac{\partial X^i}{\partial u^j} \right) \\
&= \sigma_\gamma {R_\beta}^\gamma{}_{ij} \rho^\beta X^i Y^j + \sigma_\gamma {\rho^\gamma}_{|i} \big(X(Y^i) - Y(X^i) \big) \\
&= \sigma_\gamma {R_\beta}^\gamma{}_{ij} \rho^\beta X^i Y^j + \sigma_\gamma {\rho^\gamma}_{|i} \, [X, Y]^i \\
&= \sigma_\gamma {R_\beta}^\gamma{}_{ij} \rho^\beta X^i Y^j + D_{[X,Y]} \rho
\end{aligned}$$

$$\big(D_X D_Y - D_Y D_X - D_{[X,Y]} \big) \rho = \sigma_\gamma {R_\beta}^\gamma{}_{ij} \rho^\beta X^i Y^j$$

where we have set, as before

$${R_\beta}^\gamma{}_{ij} = \frac{\partial \Gamma^\gamma_{\beta j}}{\partial u^i} - \frac{\partial \Gamma^\gamma_{\beta i}}{\partial u^j} + \Gamma^\gamma_{\alpha i} \Gamma^\alpha_{\beta j} - \Gamma^\gamma_{\alpha j} \Gamma^\alpha_{\beta i}$$

This expression $D_X D_Y - D_Y D_X - D_{[X,Y]}$ is one way to define the curvature tensor of the bundle. We have shown that the result is the same as using the other possible definition $d\omega + \omega \wedge \omega$. The equivalence of the two has been shown by showing they both give the same expression $R_{\beta}{}^{\gamma}{}_{ij}$ which is a little inelegant but does the job.

10. THE CURVATURE FORM AS AN ENDOMORPHISM OF THE SPACE OF SECTIONS

Let us reexamine the curvature form from a new point of view. Recall

$$\begin{aligned} R_{\beta}{}^{\alpha}{}_{ij} &= \frac{\partial \Gamma^{\alpha}_{\beta j}}{\partial u^i} - \frac{\partial \Gamma^{\alpha}_{\beta i}}{\partial u^j} + \Gamma^{\alpha}_{\gamma i}\Gamma^{\gamma}_{\beta j} - \Gamma^{\alpha}_{\gamma j}\Gamma^{\gamma}_{\beta i} \\ \Omega_{\beta}{}^{\alpha} &= \frac{1}{2} R_{\beta}{}^{\alpha}{}_{ij} du^i \wedge du^j = \sum_{i<j} R_{\beta}{}^{\alpha}{}_{ij} du^i \wedge du^j \\ \Omega(\rho) &= \sigma_\alpha \Omega_{\beta}{}^{\alpha} \rho^{\beta} \end{aligned}$$

Note that $R_{\beta}{}^{\alpha}{}_{ji} = -R_{\beta}{}^{\alpha}{}_{ij}$ which is a trivial but very important fact. If we input $X, Y \in T(M)$ we have (locally)

$$\begin{aligned} \Omega_{\beta}{}^{\alpha}(X,Y) &= \frac{1}{2} R_{\beta}{}^{\alpha}{}_{ij} du^i \wedge du^j (X,Y) \\ &= \frac{1}{2} R_{\beta}{}^{\alpha}{}_{ij} \left(du^i(X) du^j(Y) - du^j(X) du^i(Y) \right) \end{aligned}$$

It is very important to understand this action of $du^i \wedge du^j$ on a pair (X,Y) of tangent vectors which is

$$(du^i \wedge du^j)(X,Y) = du^i(X)du^j(Y) - du^j(X)du^i(Y)$$

Continuing

$$\begin{aligned} \Omega_{\beta}{}^{\alpha}(X,Y) &= \frac{1}{2} R_{\beta}{}^{\alpha}{}_{ij} (X^i Y^j - X^j Y^i) \\ &= \frac{1}{2} \left(R_{\beta}{}^{\alpha}{}_{ij} X^i Y^j - R_{\beta}{}^{\alpha}{}_{ij} X^j Y^i \right) \\ &= \frac{1}{2} \left(R_{\beta}{}^{\alpha}{}_{ij} X^i Y^j + R_{\beta}{}^{\alpha}{}_{ji} X^j Y^i \right) \\ &= R_{\beta}{}^{\alpha}{}_{ij} X^i Y^j \end{aligned}$$

Note how and why the $\frac{1}{2}$ disappears when we input vectors into the form $\Omega_{\beta}{}^{\alpha}$. Note that $\Omega_{\beta}{}^{\alpha}(X,Y)$ is now a matrix of $n \times n$ numbers. Thus it is suitable for describing an endomorphism of the space of sections $\Gamma(\mathcal{E})$. Locally we have

$\rho = \sigma_\alpha \rho^\alpha$ (σ_α a local basis for the sections, ρ a section). Then the coordinates are $\rho^1, \ldots, \rho^n$ and the coordinates of $\Omega(X,Y)(\rho)$ are

$$\Omega_\beta{}^1(X,Y)\rho^\beta, \ldots, \Omega_\beta{}^n(X,Y)\rho^\beta$$

and thus (in local description)

$$\Omega(X,Y)(\rho) = \sigma_\alpha \Omega_\beta{}^\alpha(X,Y)\rho^\beta$$

Def $\Omega(X,Y)$ is the *Curvature Endomorphism*

$$\Omega(X,Y) \in \mathrm{Hom}(\Gamma(\mathcal{E}), \Gamma(\mathcal{E})) \approx \Gamma(\mathcal{E}) \otimes \Gamma(\mathcal{E})^*$$

and thus

$$\Omega \in \Gamma(\mathcal{E}) \otimes \Gamma(\mathcal{E})^* \otimes \Lambda^2(T(M)^*)$$

If we input two tangent vectors into Ω it outputs an endomorphism of $\Gamma(\mathcal{E})$

Note that $R_\beta{}^\alpha{}_{ji} = -R_\beta{}^\alpha{}_{ij}$ implies that

$$\Omega(Y,X) = -\Omega(X,Y)$$

11. HIGHER COVARIANT DERIVATIVES

We want to discuss higher covariant derivatives in this section. We can simplify the computations by introducing another notation for connections. Let $\sigma_1, \cdots, \sigma_n$ be a local basis of sections of the vector bundle $\mathcal{E}$ and let $e_1 = \frac{\partial}{\partial u_1}, \cdots, e_d = \frac{\partial}{\partial u_d}$ be the natural basis of the sections of $T(M)$. Recall that $D_X\sigma_\alpha = D\sigma_\alpha(X)$ so we can rewrite

$$D_{e_i}\sigma_\alpha = \sigma_\beta \Gamma^\beta_{\alpha i}$$

as

$$\begin{aligned} D\sigma_\alpha(e_i) &= \sigma_\beta \Gamma^\beta_{\alpha i} \\ &= \sigma_\beta \Gamma^\beta_{\alpha j} du^j(e_i) \end{aligned}$$

and thus we can write

$$\begin{aligned} D\sigma_\alpha &= \sigma_\beta \Gamma^\beta_{\alpha j} du^j \\ &= \sigma_\beta\ \omega_\alpha{}^\beta \end{aligned}$$

where we put

$$\omega_\alpha{}^\beta = \Gamma^\beta_{\alpha j} du^j$$

Hence we can specify a connection by giving the 1-forms $\omega_\alpha{}^\beta$ instead of the $\Gamma^\beta_{\alpha j}$. The $\omega_\alpha{}^\beta$ are quite convenient for many types of calculations.

We now define vector valued p-forms as elements of

$$\Gamma(\mathcal{E}) \otimes \Gamma\big(\Lambda^p(T(M))\big)$$

and extend the operator D so that it acts on these spaces by Leibniz' Rule.

$$\begin{aligned} D : \Gamma(\mathcal{E}) \otimes \Gamma\big(\Lambda^p(T(M))\big) &\to \Gamma(\mathcal{E}) \otimes \Gamma\big(\Lambda^{p+1}(T(M))\big) \\ D(\rho \otimes \eta) &= \rho \otimes d\eta + D\rho \otimes \eta \end{aligned}$$

where $\rho \in \Gamma(\mathcal{E})$ and $\eta \in \Gamma\big(\Lambda^p(T(M))\big)$. We just saw that $D\sigma_\alpha = \sigma_\beta \omega_\alpha^{\ \beta}$ so we have

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

Applying the rule again we have

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

where we have used the curvature 2-form

$$\Omega_\beta^{\ \alpha} = d\omega_\beta^{\ \alpha} + \omega_\gamma^{\ \alpha} \wedge \omega_\beta^{\ \gamma}$$

We will also need to know $d\Omega$. First we note that

$$d\omega_\beta^{\ \alpha} = \Omega_\beta^{\ \alpha} - \omega_\gamma^{\ \alpha} \wedge \omega_\beta^{\ \gamma}$$

Now we compute

$$\begin{aligned} d\Omega_\beta^{\ \alpha} &= d(d\omega_\beta^{\ \alpha} + \omega_\gamma^{\ \alpha} \wedge \omega_\beta^{\ \gamma}) \\ &= 0 + d\omega_\gamma^{\ \alpha} \wedge \omega_\beta^{\ \gamma} - \omega_\gamma^{\ \alpha} \wedge d\omega_\beta^{\ \gamma} \\ &= (\Omega_\gamma^{\ \alpha} - \omega_\rho^{\ \alpha} \wedge \omega_\gamma^{\ \rho}) \wedge \omega_\beta^{\ \gamma} - \omega_\gamma^{\ \alpha} \wedge (\Omega_\beta^{\ \gamma} - \omega_\rho^{\ \gamma} \wedge \omega_\beta^{\ \rho}) \\ &= \Omega_\gamma^{\ \alpha} \wedge \omega_\beta^{\ \gamma} - \omega_\gamma^{\ \alpha} \wedge \Omega_\beta^{\ \gamma} \end{aligned}$$

which we can write as

$$d\Omega = [\Omega, \omega]$$

The presence of $[\Omega, \omega]$ gives this a faint whiff of Lie Algebra. We will see later that indeed this is the case, but need to take another approach for this to become clearer.

12. MORE NOTATION; A NEW $\mathcal{D}$

Before we proceed to the next complex of material it is convenient to introduce some new notation. Recall the equation for the Covariant Differential of a section $\rho = \sigma_\beta \rho^\beta$:

$$D\rho = D(\sigma_\beta \rho^\beta) = \sigma_\alpha (d\rho^\alpha + \omega_\beta{}^\alpha \rho^\beta)$$

where σ_α is a local basis of sections and $\omega_\beta{}^\alpha$ the corresponding expression for the connection. One might wonder, if the σ_α are going to remain fixed for the duration of an argument, why we need to think about them at all? Why not just work in $\mathbb{R}^n$ (or $\mathbb{C}^n$)? This is perfectly reasonable but causes notational stress. D is an operator on sections; we need a corresponding operator on $\mathbb{R}^n$ (or $\mathbb{C}^n$) which mirrors the activity of D. It is natural (and standard) to call this operator D also, but for awhile we will distinguish the two. Thus we want an operator $\mathcal{D}$ satisfying

$$D\rho = (\sigma_1, \ldots, \sigma_n)\mathcal{D} \begin{pmatrix} \rho^1 \\ \vdots \\ \rho^n \end{pmatrix}$$

Since also

$$D\rho = D(\sigma_\alpha \rho^\alpha) = \sigma_\alpha (d\rho^\alpha + \omega_\beta{}^\alpha \rho^\alpha)$$

we have

$$\mathcal{D} \begin{pmatrix} \rho^1 \\ \vdots \\ \rho^n \end{pmatrix} = \begin{pmatrix} d\rho^1 + \omega_\beta{}^1 \rho^\beta \\ \vdots \\ d\rho^n + \omega_\beta{}^n \rho^\beta \end{pmatrix}$$

and so

$$\mathcal{D} \begin{pmatrix} \rho^1 \\ \vdots \\ \rho^n \end{pmatrix} = (d + \omega) \begin{pmatrix} \rho^1 \\ \vdots \\ \rho^n \end{pmatrix}$$

and so we can write

$$\boxed{\mathcal{D} = d + \omega}$$

In coordinates this is

$$\mathcal{D}\rho^\alpha = d\rho^\alpha + \omega_\beta{}^\alpha \rho^\beta$$

Once again, the difference between D and $\mathcal{D}$ is that D operates on sections and $\mathcal{D}$ on the corresponding n-tuples. Any fact about one has a corresponding fact about the other; it is almost true that $\mathcal{D}$ is the matrix representation of D. Also note that the notation $\mathcal{D}\rho^\alpha$ makes it look like $\mathcal{D}$ is operating on the single section ρ^α; this is emphatically not the case; $\mathcal{D}\rho^\alpha$ depends on the entire column $(\rho^1, \ldots, \rho^n)^\top$. If we keep this in mind, however, the form $\mathcal{D}\rho^\alpha = d\rho^\alpha + \omega_\beta{}^\alpha \rho^\beta$ is handy for computation.

We now want to duplicate the calculations from the previous section in this new notation; that is to find $\mathcal{D}^2$ We will do this in two ways; first by working on the columns themselves and then by use of indices. Notice in the following that at one point the d hops a one form and this explains the minus sign.

$$\begin{aligned}
\mathcal{D}^2 \begin{pmatrix} \rho^1 \\ \vdots \\ \rho^n \end{pmatrix} &= (d+\omega)\wedge(d+\omega) \begin{pmatrix} \rho^1 \\ \vdots \\ \rho^n \end{pmatrix} \\
&= (d+\omega)\wedge\Big(\begin{pmatrix} d\rho^1 \\ \vdots \\ d\rho^n \end{pmatrix} + \omega \begin{pmatrix} \rho^1 \\ \vdots \\ \rho^n \end{pmatrix} \Big) \\
&= d\omega \begin{pmatrix} \rho^1 \\ \vdots \\ \rho^n \end{pmatrix} - \omega \wedge \begin{pmatrix} d\rho^1 \\ \vdots \\ d\rho^n \end{pmatrix} + \omega \wedge \begin{pmatrix} d\rho^1 \\ \vdots \\ d\rho^n \end{pmatrix} \\
&\quad + \omega \wedge \omega \begin{pmatrix} \rho^1 \\ \vdots \\ \rho^n \end{pmatrix} \\
&= (d\omega + \omega \wedge \omega) \begin{pmatrix} \rho^1 \\ \vdots \\ \rho^n \end{pmatrix} \\
&= \Omega \begin{pmatrix} \rho^1 \\ \vdots \\ \rho^n \end{pmatrix}
\end{aligned}$$

where of course Ω is our old friend the curvature 2-form. Now we wish to perform the same calculations componentwise. The steps match exactly those above, but look more friendly.

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

Now we must look at what happens to $\mathcal{D}$ when we change the basis. For each choice of σ we have a corresponding $\mathcal{D}_\sigma$ so that

$$D\rho = \sigma_\alpha \mathcal{D}_\sigma \rho^\alpha \qquad\qquad D\rho = \tilde{\sigma}_\alpha \mathcal{D}_{\tilde{\sigma}} \tilde{\rho}^\alpha$$

How do $\mathcal{D}_\sigma$ and $\mathcal{D}_{\tilde{\sigma}}$ relate? (Remember, although we write for ease of presentation $\mathcal{D}\rho^\alpha$, $\mathcal{D}$ actually operates on the column vector $(\rho^1, \dots, \rho^n)^\top$.) We have

$$\tilde{\sigma}_\alpha = \sigma_\beta\, g^\beta_\alpha \quad \text{and} \quad \sigma_\alpha = \tilde{\sigma}_\beta\, h^\beta_\alpha$$

where (g_α^β) and (h_β^α) are inverse matrices from GL(n) or some subgroup. Thus

$$\begin{aligned} \sigma_\gamma \rho^\gamma &= \rho = \tilde{\sigma}_\beta \tilde{\rho}^\beta \\ &= \sigma_\gamma g_\beta^\gamma \tilde{\rho}^\beta \end{aligned}$$

so

$$\rho^\gamma = g_\beta^\gamma \tilde{\rho}^\beta$$

Thus

$$\begin{aligned} \tilde{\sigma}_\alpha \mathcal{D}_{\tilde{\sigma}} \tilde{\rho}^\alpha = D\rho &= \sigma_\alpha \mathcal{D}_\sigma \rho^\alpha \\ \tilde{\sigma}_\alpha (d\tilde{\rho}^\alpha + \tilde{\omega}_\beta{}^\alpha d\tilde{\rho}^\beta) &= \sigma_\gamma (d\rho^\gamma + \omega_\delta{}^\gamma \rho^\delta) \\ &= \tilde{\sigma}_\alpha h_\gamma^\alpha \big(d(g_\beta^\gamma \tilde{\rho}^\beta) + \omega_\delta{}^\gamma g_\beta^\delta \tilde{\rho}^\beta \big) \\ &= \tilde{\sigma}_\alpha \big(h_\gamma^\alpha g_\beta^\gamma d\tilde{\rho}^\beta + (h_\gamma^\alpha dg_\beta^\gamma) \tilde{\rho}^\beta + h_\gamma^\alpha \omega_\delta{}^\gamma g_\beta^\delta \tilde{\rho}^\beta \big) \\ &= \tilde{\sigma}_\alpha \big(d\tilde{\rho}^\alpha + (h_\delta^\alpha dg_\beta^\delta + h_\gamma^\alpha \omega_\delta{}^\gamma g_\beta^\delta) \tilde{\rho}^\beta \big) \end{aligned}$$

Comparing the two sides we see that the requirement that $D\rho$ be independent of basis change requires that the connection coefficients $\omega_\delta{}^\gamma$ for σ and $\tilde{\omega}_\beta{}^\alpha$ for $\tilde{\sigma}$ be connected by the equation

$$\tilde{\omega}_\beta{}^\alpha = h_\gamma^\alpha \omega_\delta{}^\gamma g_\beta^\delta + h_\delta^\alpha \, dg_\beta^\delta$$

or in matrix notation

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

From the standpoint of Lie Groups and Lie Algebras these formulas are familiar; $\omega \in \mathfrak{g}l(n, \mathbb{R}) = M(n, \mathbb{R})$, $g \in \mathrm{GL}(n, \mathbb{R})$ and we have

$$\tilde{\omega} = \omega_G + \mathrm{Ad}(g^{-1})\omega$$

where $\omega_G = g^{-1} dg$ is the canonical left invariant form on the Lie Group G and Ad is the left regular representation of the Lie Group on the Lie Algebra as described in the next section.

13. SOME ABSTRACT CONSIDERATIONS

Before going on to other aspects of connections let us pause for a moment to consider the formula

$$\tilde{\omega} = \omega_G + \mathrm{Ad}(g^{-1})\omega$$

Here $\omega_G = g^{-1} dg$ is the canonical left invariant form on G and Ad is the left regular representation of G on the Lie Algebra $\mathfrak{g} = \mathrm{M}(n, \mathbb{R})$.

As before let $\sigma = (\sigma_1, \dots, \sigma_n)$ and and $\tilde{\sigma} = (\tilde{\sigma_1}, \dots, \tilde{\sigma}_n)$ be local bases of sections over the open set $U \subseteq M$ which will usually be a coordinate neighborhood. Then we have

$$\tilde{\sigma} = \sigma g \qquad\qquad \tilde{\sigma}_\beta = \sigma_\alpha g_\beta^\alpha$$

where $g = (g^\alpha_\beta)$ is a matrix valued function defined over U. We can form a new bundle $\mathcal{F}$ from $\mathcal{E}$ whose fibre over $x \in M$ consists of all bases of $\pi^{-1}(x)$, and then σ and $\tilde{\sigma}$ are sections of $\mathcal{F}$. $\mathcal{F}$ is called the *frame bundle* for $\mathcal{E}$; it is of course not a vector bundle. Given some particular section σ of $\mathcal{F}$ any matrix function $g : U \to G = \mathrm{GL}(n, \mathbb{R})$ produces a new section $\tilde{\sigma}$ of $\mathcal{F}$ via $\tilde{\sigma} = \sigma g$, and all sections of $\mathcal{F}$ arise in this way. Thus there is a bijection between functions $g : U \to G = \mathrm{GL}(n, \mathbb{R})$ and sections of $\mathcal{F}$. On each fibre, $G = \mathrm{GL}(n, \mathbb{R})$ moves us around freely and transitively in the fibre, and this is consistent with the group structure. For let $\tilde{\sigma} = \sigma g_1$ and $\tilde{\tilde{\sigma}} = \tilde{\sigma} g_2$. Then

$$\tilde{\tilde{\sigma}} = \tilde{\sigma} g_2 = (\sigma g_1) g_2 = \sigma(g_1 g_2)$$

so we are dealing with a RIGHT G ACTION on the fibres of $\mathcal{F}$.

This gives us the option of replacing the fibres of $\mathcal{F}$ with the group G itself, once we have chosen a "base section" σ_0 to use in correlating each $\sigma \in \mathcal{F}$ with it's corresponding group element $g \in G$. However, note that this correlation is not canonical. (Note we are now suppressing the constantly necessary "at each $x \in M$" in this paragraph and in what follows.)

Since we are dealing with a general vector bundle, for us the structure group is $G = \mathrm{GL}(n, \mathbb{R})$. This is the group the transition maps live in. The new structure, with the fibres all isomorphic to G, is our first example of a Principal Bundle. We will go into this in much more detail later; here we are just trying to get a glimpse of the big picture.

For more of the big picture, we are now going to recognize that the structure group $G = \mathrm{GL}(n, \mathbb{R})$ is a Lie Group. Thus it has a Lie Algebra, $\mathfrak{g} = \mathfrak{g}l(n, \mathbb{R}) = \mathrm{M}(n, \mathbb{R})$, which is simply all n by n real matrices. In subsequent work we will be concerned with Metric Bundles, vector bundles where each fibre carries an inner product, and then the structure group will be $G = \mathrm{O}(n, \mathbb{R})$, the orthogonal group, and the Lie Algebra will then be $\mathfrak{o}(n, \mathbb{R})$, the set of real skew symmetric matrices.

Now recall that

$$\omega = ({\omega_\beta}^\alpha) = (\Gamma^\alpha_{\beta i}\, du^i)$$

so that when we insert a tangent vector $X \in T_p(M)$ we get a real n by n matrix. which is in the Lie Algebra of the structure group $G = \mathrm{GL}(n, \mathbb{R})$. Hence we can say that ω is a Lie Algebra valued one form; insert tangent vector, get Lie Algebra element.

Exactly the same will occur with other structure groups; if the fibres carry inner products and the structure group is $\mathrm{O}(n, \mathbb{R})$ then $\omega(X)$ will be in $\mathfrak{o}(n, \mathbb{R})$, so it will be a skew symmetric matrix.

We now return to our case of $G = \mathrm{GL}(n, \mathbb{R})$ and we want to explain how the change of basis formula for the connection ω relates to the Lie Group. To do this we must first explain the left regular representation of a matrix Lie Group G on it's Lie Algebra, which works the same for any Matrix Lie Group (and can be generalized to any Lie Group though not with these formulas).

We define the action of $g \in G$ on an element $A \in \mathfrak{g}$ by

$$g \cdot A = gAg^{-1}$$

This is a left action since

$$\begin{aligned} g_1 \cdot (g_2 \cdot A) &= g_1(g_2 \cdot A)g_1^{-1} = g_1 g_2 A g_2^{-1} g_1^{-1} \\ &= g_1 g_2 A (g_1 g_2)^{-1} = (g_1 g_2) \cdot A \end{aligned}$$

so that indeed we have a left action. We now switch to the Ad notation where we replace $g \cdot A$ by $(\mathrm{Ad}(g))(A)$ which we write in a less clear but more civilized way as $\mathrm{Ad}(g)(A)$ or $\mathrm{Ad}\, g(A)$ or $\mathrm{Ad}(g)A$. In this notation the previous equation becomes

$$\mathrm{Ad}\, g_1(\mathrm{Ad}\, g_2(A)) = \mathrm{Ad}(g_1 g_2)(A)$$

or

$$\mathrm{Ad}\, g_1 \circ \mathrm{Ad}\, g_2 = \mathrm{Ad}(g_1 g_2)$$

which shows that Ad is a representation of G on its Lie Algebra $\mathfrak{g}$

We now want to show that the transformation formula for ω represents a right action of G. To this end let ω correspond to the basis σ, $\tilde{\omega}$ correspond to the basis $\tilde{\sigma} = \sigma g_1$ and $\tilde{\tilde{\omega}}$ correspond to $\tilde{\tilde{\sigma}} = \tilde{\sigma} g_2$. Recall that the transformation formulas are

$$\begin{aligned} \tilde{\omega} \,=\, \omega \cdot g_1 &= g_1^{-1}\, dg_1 + \mathrm{Ad}(g_1^{-1})\omega \\ \tilde{\tilde{\omega}} \,=\, \tilde{\omega} \cdot g_2 &= g_2^{-1}\, dg_2 + \mathrm{Ad}(g_2^{-1})\tilde{\omega} \end{aligned}$$

Thus we have

$$\begin{aligned} (\omega \cdot g_1) \cdot g_2 \,=\, \tilde{\omega} \cdot g_2 &= g_2^{-1}\, dg_2 + \mathrm{Ad}(g_2^{-1})\tilde{\omega} \\ &= g_2^{-1}\, dg_2 + \mathrm{Ad}(g_2^{-1})\big[g_1^{-1}\, dg_1 + \mathrm{Ad}(g_1^{-1})\tilde{\omega}\big] \\ &= g_2^{-1}\, dg_2 + g_2^{-1} g_1^{-1} (dg_1) g_2 + \mathrm{Ad}(g_2^{-1})\mathrm{Ad}(g_1^{-1})\omega \\ &= g_2^{-1}\big[dg_2 + g_1^{-1}(dg_1)g_2\big] + \mathrm{Ad}(g_2^{-1}g_1^{-1})\omega \\ &= g_2^{-1}g_1^{-1}\big[g_1(dg_2) + (dg_1)g_2\big] + \mathrm{Ad}(g_1 g_2)^{-1})\omega \\ &= (g_1 g_2)^{-1} d(g_1 g_2) + \mathrm{Ad}((g_1 g_2)^{-1})\omega \\ &= \omega \cdot (g_1 g_2) \end{aligned}$$

as required for a right action.

While we are dealing with Lie Algebra valued one-forms we will take the opportunity to introduce some new notation used in many sources. This is the Lie Bracket for forms. The definitions are slightly counterintuitive because of the way certain conventions interact.

We will first discuss the case where the structure group is a matrix group. The user should keep $\mathrm{GL}(n, \mathbb{R})$ in mind, although what we say will work for any other matrix group.

Def Let ω and η be one forms. Then

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

Setting $\eta = \omega$ we have

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

Hence by the standard conventions we have, for the two form $[\omega, \eta]$,

$$[\omega, \eta](X, Y) = 2\omega \wedge \eta(X, Y) = 2(\omega(X)\eta(Y) - \omega(Y)\eta(X))$$

Now watch what happens when we put $\eta = \omega$:

$$\begin{aligned} [\omega, \omega](X, Y) &= 2(\omega(X)\omega(Y) - \omega(Y)\omega(X)) \\ &= 2[\omega(X), \omega(Y)] \end{aligned}$$

Since $[\omega, \omega] = 2\omega \wedge \omega$, we can write

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

as

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

and this is used is some advanced books like Berline.

The user may well ask why we would exchange the perfectly understandable $\Omega = d\omega + \omega \wedge \omega$ for the less understandable $\Omega = d\omega + \frac{1}{2}[\omega, \omega]$. There is indeed a reason, and the reason turns on the equation $[\omega, \omega](X, Y) = 2(\omega(X)\omega(Y) - \omega(Y)\omega(X))$ The right half of this equation makes sense when we are dealing with structure groups which are matrix Lie Groups. However, if the Groups are NOT matrix Lie Groups, but more general Groups (Spinor Groups for example) then the Lie Algebra may not be a subset of a matrix group and thus the right side of the equation makes no sense. However, since we are always in a Lie Algebra, $2[\omega(X), \omega(Y)]$ DOES make sense. Thus if we simply require

$$[\omega, \omega](X, Y) = 2[\omega(X), \omega(Y)]$$

we generalize to the more abstract situation.

14. PARALLEL TRANSFER IN A VECTOR BUNDLE

One of the most important uses of connections is in the concept of Parallel Transfer. This is important not only for its uses in vector bundles themselves but also as a way to make the transition to Principal Bundles.

Let ρ_x be a vector in the fibre of $\mathcal{E}$ over x and let $x(t)$ be a curve on M with coordinates $u^1(t), \ldots, u^d(t)$ where $x(0) = x = \pi(\rho_x)$. We are interested here in the portion of the curve that lies in a coordinate patch U containing x. Let $\sigma = (\sigma_1, \ldots, \sigma_n)$ be a local basis of sections over U and $\omega = ({\omega_\alpha}^\beta)$ a connection written in terms of σ;

$$D\sigma_\alpha = \sigma_\beta \, {\omega_\alpha}^\beta$$

We wish to create a vector function $\rho(t) \in \mathcal{E}$ lying over $x(t)$ (with coordinates $u^1(t), \ldots, u^d(t)$) which coincides with ρ_x at $x(0)$ and which has all the vectors in the family "parallel" to one another. Given our environment there is pretty much only one thing that "parallel" can mean; that the directional derivative in the direction of the curve $x(t)$ of $\rho(t)$ must be 0. There is also not much doubt of how to express this. Indeed we must have

Def The section $\rho(t)$ of the bundle $\mathcal{E}$ along the curve $x(t)$ is *parallel* if and only if

$$D_{\dot{x}(t)}\rho(t) = 0$$

Thus $\rho(0)$ becomes one of a family of parallel vectors along the curve. Naturally the vector $\rho(t_1)$ does not depend on $x(t_1)$ alone; it is necessary to know *by which path* we got from $x(0)$ to $x(t_1)$. A different path $y(u)$ (with $x(0) = y(0)$ and $x(t_1) = y(u_1)$) would, in general, give a $\tilde{\rho}(y(u_1)) \neq \rho(x(t_1))$. It is very important to keep this *path dependence* in mind.

Note that parallel transport gives a *linear isomorphism* $P_t : \pi^{-1}[x(t)] \to \pi^{-1}[x(0)]$ of the various fibres along the curve $x(t)$

If E has certain additional structures, for example an inner product on the fibres, then parallel transport may respect these structures provided the connection satisfies certain additional conditions. The situation clarifies nicely in principal bundles and we will treat it there.

We now want to decode the defining equation $D_{\dot{x}(t)}\rho(t) = 0$ to see what the condition really says. This is easy.

$$\begin{aligned}
D_{\dot{x}(t)}\rho(t) &= 0 \\
D\big(\sigma_\alpha(t)\rho^\alpha(t)\big)\big(\dot{x}(t)\big) &= 0 \\
\sigma_\alpha(t)d\rho^\alpha(t)\big(\dot{x}(t)\big) + D\big(\sigma_\alpha(t)\big)\big(\dot{x}(t)\big)\rho^\alpha(t) &= 0 \\
\sigma_\beta(t)\Big[d\rho^\beta(t)\big(\dot{x}(t)\big) + \omega_\alpha{}^\beta\big(\dot{x}(t)\big)\rho^\alpha(t)\Big] &= 0
\end{aligned}$$

The $\omega_\alpha{}^\beta$ are 1-forms on $\Gamma(T(M))$ and hence can be expressed in the local basis $du^1, \dots, du^d$ dual to $\frac{\partial}{\partial u^1}, \dots, \frac{\partial}{\partial u^d}$ by

$$\omega_\alpha{}^\beta = \Gamma^\beta_{\alpha i}\, du^i$$

and $d\rho^\beta(t)(\dot{x}(t))$ decodes as

$$\frac{\partial \rho^\beta}{\partial u^j}\, du^j\big(\dot{u}^i(t)\frac{\partial}{\partial u^i}\big) = \frac{\partial \rho^\beta}{\partial u^j}\frac{du^j}{dt} = \frac{d\rho^\beta}{dt}$$

and $\omega_\alpha{}^\beta(\dot{x}(t))$ decodes as

$$\omega_\alpha{}^\beta(\dot{x}(t)) = \Gamma^\beta_{\alpha i}\, du^i\Big(\frac{du^j}{dt}\frac{\partial}{\partial u^j}\Big) = \Gamma^\beta_{\alpha i}\,\frac{du^i}{dt}$$

so we finally come to

$$\frac{d\rho^\beta}{dt} + \Gamma^\beta_{\alpha i}\,\frac{du^i}{dt}\rho^\alpha = 0$$

as the differential equations for parallel transfer along the curve $x(t)$ on M. Since $x(t)$ and thus $\dot{x}(t)$ are given, this is an autonomous system of ordinary differential equations for the $\rho^\beta(t)$ and, given the initial value $\rho^\beta(0)$, has a unique solution (at least for small t). (This is very standard stuff.)

15. ANOTHER WAY OF LOOKING AT PARALLEL TRANSFER IN A VECTOR SPACE; THE HORIZONTAL SPACE

To begin the process of moving from Vector Bundles to Principal Bundles we need to look at the vector bundle picture from a different point of view. Once again we consider parallel transfer in a vector bundle. Using coordinates $u^1, \ldots, u^d$ on M and $\rho^1, \ldots, \rho^n$ for the section $\sigma_\alpha \rho^\alpha$ (σ_α a local basis of Sections) the differential equations for parallel transfer along the curve $x(t)$ with coordinates $u^i(t)$ are

$$\frac{d\rho^\alpha}{dt} + \Gamma^\alpha_{\beta i} \frac{du^i}{dt} \rho^\beta = 0$$

with initial conditions

$$\rho^\alpha(0) = \rho^\alpha_0 \qquad \text{at} \qquad t = 0$$

Then the vectors $\rho(t) = \sigma_\alpha \rho^\alpha(t)$ form a parallel family along the curve $x(t)$.

Our new view of this situation is to contemplate a curve (also denoted by $\rho(t)$) *in the vector bundle $\mathcal{E}$ itself.* The curve "lies over" $x(t)$ in the sense that $\pi(\rho(t)) = x(t)$, and again consists of parallely transferred vectors. We can use $u^1, \ldots, u^d, \rho^1, \ldots, \rho^n$ as coordinates for the curve in $\mathcal{E}$, since locally $\mathcal{E}$ is isomorphic to $U \times \mathbb{R}^n$. The differential equations for the curve $\rho : [0, a] \to \mathcal{E}$ are

$$\begin{aligned} \frac{du^i}{dt} &= \dot{u}^i \qquad (\dot{u}^i \text{ known from the equation of the curve}) \\ \frac{d\rho^\alpha}{dt} &= -\Gamma^\alpha_{\beta i} \rho^\beta \dot{u}^i = -\omega_\beta{}^\alpha(\dot{x}(t)) \end{aligned}$$

The solution $\rho(t)$ with coordinates $(u^i(t), \rho^\alpha(t))$ is called the *lift* of $x(t)$ into the bundle with initial conditions $\rho^\alpha(0) = \rho^\alpha_0$. Thus parallel transfer is the same thing as lifting $x(t)$ into $\mathcal{E}$ with the above equations and initial conditions.

We remark that lifting curves up projections is a very common sort of mathematical activity, for example in the theory of covering spaces.

Next we want to interpret the above equations in terms of the tangent space of the bundle $\mathcal{E}$. Notice that $T(\mathcal{E})$ is a new actor on the stage. We can use $\frac{\partial}{\partial u^1}, \ldots, \frac{\partial}{\partial u^d}, \frac{\partial}{\partial \rho^1}, \ldots, \frac{\partial}{\partial \rho^n}$ as basis for $T(\mathcal{E})$.

The differential forms in $T(\mathcal{E})^*$ have a basis $du^1, \ldots, du^d, d\rho^1, \ldots, d\rho^n$. Now consider the lift of any curve $x(t)$ starting at $x(0)$ and with initial vector $\rho_0 = \sigma_\alpha \rho^\alpha_0$. The tangent vector $dx/dt|_{t=0}$ with coordinates $du^i/dt|_{t=0}$ will vary as we change the curve, but we will always have

$$\frac{d\rho^\alpha}{dt} + \Gamma^\alpha_{\beta i} \rho^\beta \frac{du^i}{dt} = 0$$

for all the lifts. Hence the tangent vectors to all such curves will satisfy

$$d\rho^\alpha + \Gamma^\alpha_{\beta i} \rho^\beta du^i = 0$$

which is a system of n linear conditions in $T_{\rho(0)}(\mathcal{E})$ and will thus lie in an $n + d - n = d$ dimensional subspace of the tangent space $T(\mathcal{E})$ at (u_0, ρ_0) This subspace is the famous *Horizontal Subspace* $H(\mathcal{E})$ of $T(\mathcal{E})$. The *Vertical Subspace* $V(\mathcal{E})$ is the subspace of $T(\mathcal{E})$ of vectors tangent to the fibre, and thus is n-dimensional.It is defined by the equations $du^i = 0$.

Note that if $v \in V(\mathcal{E}) \cap H(\mathcal{E})$ then $du^i(v) = 0$ and since $d\rho^\alpha + \Gamma^\alpha_{\beta i}\rho^\beta du^i = 0$ we then have $d\rho^\alpha(v) = 0$ and so $v = 0$.

Since $T(\mathcal{E})$ is $n + d$ dimensional, we then have

$$T(\mathcal{E}) = V(\mathcal{E}) \oplus H(\mathcal{E})$$

The collection of all $H(\mathcal{E})$ for each $x \in M$ forms the *Horizontal Subbundle* (also denoted by $H(\mathcal{E})$).

Note that $H(\mathcal{E})$ is defined only through the use of the connection 1-form $\omega_\beta{}^\alpha$. *There is no horizontal bundle without a connection.* Indeed, from the most advanced view a connection *is* a specification of the Horizontal Bundle $H(\mathcal{E})$ at each point of $\mathcal{E}$.

It is unusual to consider the horizontal space of a vector bundle; it is more usual to consider it for a principal bundle, and we will do this. However, horizontal spaces can be defined for any sort of bundle so I thought it would be fun to practise on vector bundles first, where the ideas are a little simpler to implement and we can write down explicit equations.

Now let us recall the operator $\mathcal{D}$; we have

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

Since the σ_α are a local basis of sections of the vector bundle, we can regard $D\rho = \sigma_\alpha \mathcal{D}\rho^\alpha$ as a map from $T(\mathcal{E})$ to $V(\mathcal{E})$.

Here we are using the fact that the fibres are vector spaces and hence can be identified with their own tangent spaces and with this identification basis vectors σ^α are identified with basis vectors $\frac{\partial}{\partial \rho^\alpha}$ of the tangent space.

Let's look at what $D\rho = \sigma_\alpha \mathcal{D}\rho^\alpha$ does to a tangent vector $v \in T(\mathcal{E})$. We can write v as

$$v = v^j \frac{\partial}{\partial u^j} + V^\beta \frac{\partial}{\partial \rho^\beta}$$

so that (v^i, V^α) become coordinates of v. Then

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

Suppose now v is a vertical tangent vector; $v \in V(\mathcal{E})$. Then all $v^j = 0$ and

$$D\rho(v) = \sigma_\alpha \mathcal{D}\rho^\alpha(v) = \sigma_\alpha(V^\alpha + 0) = v$$

and thus $D\rho = \sigma_\alpha \mathcal{D}\rho^\alpha$ is the identity on $V(\mathcal{E})$. (Remember the identification $\sigma_\alpha V^\alpha = V^\alpha \frac{\partial}{\partial \rho^\alpha} = v$.) Hence $D\rho = \sigma_\alpha \mathcal{D}\rho^\alpha$ is a *projection* onto $V(\mathcal{E})$. The kernel of this projection is of course the horizontal space $H(\mathcal{E})$ since this is defined by $D\rho(v) = 0$. We often say that $D\rho$ projects $T(\mathcal{E})$ *onto* $V(\mathcal{E})$ *along* $H(\mathcal{E})$.

Conversely, suppose Π is such a projection. Then $\Pi\, v = v$ for $v \in V(\mathcal{E})$. Π can be written in terms of the basis $\frac{\partial}{\partial u^i}, \frac{\partial}{\partial \rho^\alpha}$ and thus

$$\begin{aligned} \Pi(v) &= \Pi\Big(v_i \frac{\partial}{\partial u^i} + V^\alpha \frac{\partial}{\partial \rho^\alpha}\Big) \\ &= \sigma_\alpha A^\alpha \Big(v_i \frac{\partial}{\partial u^i} + V^\alpha \frac{\partial}{\partial \rho^\alpha}\Big) \end{aligned}$$

for certain 1-forms $A^\alpha : T(\mathcal{E}) \to \mathbb{R}$. But for $v \in V(\mathcal{E})$, we have $\Pi\, v = v$, so

$$\begin{aligned} \sigma_\alpha V^\alpha &= v \;=\; \Pi\, v \\ &= \sigma_\alpha A^\alpha \Big(0 + V^\beta \frac{\partial}{\partial \rho^\beta}\Big) \\ V^\alpha &= A^\alpha \Big(V^\beta \frac{\partial}{\partial \rho^\beta}\Big) \end{aligned}$$

so

$$A^\alpha = d\rho^\alpha + (\text{terms in } du^i)$$

since the du^i terms die on $\frac{\partial}{\partial \rho^\beta}$. Linearity then forces

$$A^\alpha = d\rho^\alpha + \Gamma^\alpha_{\beta i} \rho^\beta du^i$$

for some $\Gamma^\alpha_{\beta i}$ which may be chosen arbitrarily.

Thus we have come full circle. We can define a connection relative to a local basis σ_α by a system $\Gamma^\alpha_{\beta i}\, du^i$ of 1-forms or abstractly by the requirement that it be linear and a projection from $T(\mathcal{E})$ to $V(\mathcal{E})$. In either case we get

$$\mathcal{D}\rho^\alpha = d\rho^\alpha + \omega_\beta{}^\alpha \rho^\beta = d\rho^\alpha + \Gamma^\alpha_{\beta i} \rho^\beta\, du^i$$

where the $\Gamma^\alpha_{\beta i}$ are arbitrarily chosen smooth functions which behave properly under change of local basis of sections.

Recall that the Horizontal Space $H(\mathcal{E}) = \ker D\rho$. This is a coordinate independent condition. In coordinates we have the the equations:

$$\begin{aligned} \mathcal{D}_\sigma \rho^\alpha &= \omega_\beta{}^\alpha \rho^\beta \;=\; 0 \qquad \text{basis } \sigma \\ \mathcal{D}_{\tilde\sigma} \tilde\rho^\alpha &= \tilde\omega_\beta{}^\alpha \tilde\rho^\beta \;=\; 0 \qquad \text{basis } \tilde\sigma \end{aligned}$$

These are connected to one another, with $\tilde\sigma = \sigma\, g$, by

$$\tilde\omega = g^{-1}\, dg + g^{-1} \omega g = \omega_0 + \operatorname{Ad} g^{-1}(\omega)$$

as we previously saw.

It might possibly be interesting to look at the matrix of the projection Π of $T(\mathcal{E})$ into $V(\mathcal{E})$. We set $v \in T(\mathcal{E})$ and $w \in V(\mathcal{E})$ where $w = D\rho(v) = \Pi(v)$. In coordinates we have

$$\begin{aligned} v &= v^i \frac{\partial}{\partial u^i} + V^\alpha \frac{\partial}{\partial \rho^\alpha} \\ w &= \Pi v \;=\; W^\alpha \frac{\partial}{\partial \rho^\alpha} \end{aligned}$$

where

$$W^\alpha = V^\alpha + \Gamma^\alpha_{\beta\, i}\, \rho^\beta v^i$$

Then the $n \times (n+d)$ matrix for Π is

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

The identity matrix on the right side of our matrix reflects the fact that $\Pi v = v$ for $v \in V(\mathcal{E})$.

16. THE HORIZONTAL SPACE IS ALL YOU NEED

This section introduces the third and most sophisticated method of defining curvature. It is highly geometric and readily generalizable to any fibre bundle and depends only on the Horizontal Space. However, it lacks immediate intuitive content.

As we saw in the previous section, knowledge of the Horizontal space determines the connection 1-forms and we will show in this section using lifts of vector fields that we can get directly to curvature.

The calculation in this section is not an optimal method of getting to the result, but I think it is interesting that the result can be obtained in such a completely elementary manner at this point. Perhaps this is a calculation where it might be best to skim. Part of the reason I do it is because I may have use for some of the formulas later.

Let X and Y be vector fields on M and let $\tilde{X}$ and $\tilde{Y}$ be the lifts of these vector fields to the Horizontal Space. This is defined as follows. Like $T(M)$ the horizontal space has dimension d at each point. Using the π that goes with the vector bundle, $\pi : \mathcal{E} \to M$, we see that π_* restricted to the horizontal space must be an isomorphism (at each point) and thus we can form $\tilde{X} = \pi_*^{-1} X$. It is easy to find the equations for $\tilde{X}$. Since $\tilde{X} \in H(\mathcal{E})$ it can be written in coordinates as

$$\tilde{X} = X^i \frac{\partial}{\partial u^i} + x^\alpha \frac{\partial}{\partial \rho^\alpha} \qquad \text{where} \qquad X = X^i \frac{\partial}{\partial u^i}$$

We recall that the Horizontal Space is defined by the condition

$$\Pi(\tilde{X}) = x^\alpha + \Gamma^\alpha_{\beta i} \rho^\beta X^i = 0$$

so

$$x^\alpha = -\Gamma^\alpha_{\beta i} \rho^\beta X^i$$

Thus

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

Since we have this formula, in a sense, the lifting process is trivial and so are the computations, though long and intricate.

Recall that Ω^β_α is the curvature two form from previous sections. We consider the lifts of X, Y, and $[X, Y]$ and we have the formula

Theorem $\qquad \widetilde{[X,Y]} - [\tilde{X}, \tilde{Y}] = \Omega_\delta{}^\beta(X,Y)\rho^\delta \frac{\partial}{\partial \rho^\beta}$

Thus knowledge of the Horizontal Space means knowledge of the curvature, and this is probably the most fundamental fact about curvature. Since in our circumstance the horizontal space is determined by the Christoffel symbols $\Gamma^\alpha_{\beta i}$ and so is the curvature, the result is perhaps not surprising. What is surprising is that the relationship is so simple.

I suspect the result is connected somehow to the the relation that measures the failure of the Covariant Derivative D_X and D_y to commute.

To prove the theorem we have only to compute all the quantities, which is not much fun but not particularly hard. The difficult one is $[\tilde{X}, \tilde{Y}]$. We have

$$\tilde{X} = X^i \frac{\partial}{\partial u^i} - \Gamma^\alpha_{\beta i}\rho^\beta X^i \frac{\partial}{\partial \rho^\alpha} = X^i \frac{\partial}{\partial u^i} + x^\alpha \frac{\partial}{\partial \rho^\alpha}$$
$$\tilde{Y} = Y^j \frac{\partial}{\partial u^j} - \Gamma^\gamma_{\delta i}\rho^\delta Y^j \frac{\partial}{\partial \rho^\gamma} = Y^j \frac{\partial}{\partial u^j} + y^\beta \frac{\partial}{\partial \rho^\beta}$$

We need $[\tilde{X}, \tilde{Y}]$. To get this we first compute $\tilde{X}(Y^j)$ and $\tilde{X}(y^\alpha)$.

$$\begin{aligned}
\tilde{X}(Y^j) &= X^i \frac{\partial Y^j}{\partial u^i} - \Gamma^\alpha_{\beta k} X^i \rho^\beta \frac{\partial Y^j}{\partial \rho^\alpha} \\
&= X^i \frac{\partial Y^j}{\partial u^i} - 0 \\
&= X(Y^j) \\
\tilde{X}(y^\beta) &= \tilde{X}(-\Gamma^\beta_{\delta j} Y^j \rho^\delta) \\
&= X^i \frac{\partial}{\partial u^i}\left(-\Gamma^\beta_{\delta j} Y^j \rho^\delta\right) - \Gamma^\alpha_{\gamma \ell} X^\ell \rho^\gamma \frac{\partial}{\partial \rho^\alpha}\left(-\Gamma^\beta_{\delta j} Y^j \rho^\delta\right) \\
&= -X^i \frac{\partial \Gamma^\beta_{\gamma j}}{\partial u^i} Y^j \rho^\delta - X^i \Gamma^\beta_{\delta j} \frac{\partial Y^j}{\partial u^i} \rho^\delta + \Gamma^\alpha_{\gamma \ell}\Gamma^\beta_{\alpha j} X^\ell Y^j \rho^\gamma
\end{aligned}$$

We can now already smell victory since the various pieces of ${\Omega_\delta}^\beta$ are already visible here. For $[\tilde{X}, \tilde{Y}]$ we need

$$\begin{aligned}
\tilde{X}(Y^i)\frac{\partial}{\partial u^i} + \tilde{X}(y^\beta)\frac{\partial}{\partial \rho^\beta} &= \tilde{X}(Y^i)\frac{\partial}{\partial u^i} - \tilde{X}(\Gamma^\beta_{\delta j} Y^j \rho^\delta)\frac{\partial}{\partial \rho^\beta} \\
&= X(Y^j)\frac{\partial}{\partial u^j} - \Gamma^\beta_{\delta j} X(Y^j)\rho^\delta \frac{\partial}{\partial \rho^\beta} \\
&+ \left[\left(-\frac{\partial \Gamma^\beta_{\delta j}}{\partial u^i} + \Gamma^\alpha_{\delta i}\Gamma^\beta_{\alpha j}\right) X^i Y^j \rho^\delta\right] \frac{\partial}{\partial \rho^\beta}
\end{aligned}$$

By symmetry

$$\begin{aligned}
\tilde{Y}(X^i)\frac{\partial}{\partial u^i} + \tilde{Y}(x^\alpha)\frac{\partial}{\partial \rho^\alpha} &= Y(X^j)\frac{\partial}{\partial u^j} - \Gamma^\beta_{\delta j} Y(X^j)\rho^\delta \frac{\partial}{\partial \rho^\beta} \\
&+ \left[\left(-\frac{\partial \Gamma^\beta_{\delta j}}{\partial u^i} + \Gamma^\alpha_{\delta i}\Gamma^\beta_{\alpha j}\right) Y^i X^j \rho^\delta\right] \frac{\partial}{\partial \rho^\beta}
\end{aligned}$$

Putting this together we have

$$
\begin{aligned}
[\tilde{X},\tilde{Y}] &= \big(X(Y^j)-Y(X^j)\big)\frac{\partial}{\partial u^i} - \Gamma^\beta_{\delta j}[X,Y]^j\rho^\delta\frac{\partial}{\partial\rho^\beta} \\
&+ \left[\left(-\frac{\partial\Gamma^\beta_{\delta j}}{\partial u^i}+\frac{\partial\Gamma^\beta_{\delta i}}{\partial u^j}+\Gamma^\alpha_{\delta i}\Gamma^\beta\alpha j-\Gamma^\alpha_{\delta j}\Gamma^\beta_{\alpha i}\right)X^iY^j\rho^\delta\right]\frac{\partial}{\partial\rho^\beta} \\
&= [X,Y]^j\frac{\partial}{\partial u^j} - \Gamma^\beta_{\delta j}[X,Y]^j\rho^\delta\frac{\partial}{\partial\rho^\beta} - R_\delta{}^\beta{}_{ij}X^iY^j\rho^\delta\frac{\partial}{\partial\rho^\beta} \\
&= [X,Y]^j\frac{\partial}{\partial u^j} - \Gamma^\beta_{\delta j}[X,Y]^j\rho^\delta\frac{\partial}{\partial\rho^\beta} - \Omega_\delta{}^\beta(X,Y)\rho^\delta\frac{\partial}{\partial\rho^\beta}
\end{aligned}
$$

On the other hand, using the same formula as before,

$$
\begin{aligned}
[X,Y] &= [X,Y]^j\frac{\partial}{\partial u^j} \\
\widetilde{[X,Y]} &= [X,Y]^j\frac{\partial}{\partial u^j} - \Gamma^\beta_{\delta j}[X,Y]^j\rho^\delta\frac{\partial}{\partial\rho^\beta}
\end{aligned}
$$

Thus

$$
\widetilde{[X,Y]} - [\tilde{X},\tilde{Y}] = \Omega_\delta{}^\beta(X,Y)\rho^\delta\frac{\partial}{\partial\rho^\beta}
$$

as desired

17. MORE ON THE INVARIANCE OF Π

We have seen that $\Pi = D\rho$ is described invariantly and thus is independent of the coordinate system. In spite of this there is some interest in a low level derivation of this fact, since there are some slightly tricky aspects which we might like to be aware of later. This also clarifies the role of $\omega_G = g^{-1}\,dg$ in the formula

$$\tilde{\omega} = \omega_G + \mathrm{Ad}(g^{-1})\omega$$

Recall as usual

$$\tilde{\sigma} = \sigma g, \qquad h = g^{-1}, \qquad \tilde{\rho}^\gamma = h^\gamma_\nu\rho^\nu$$

We need two preliminary results. Since $g^\alpha_\beta h^\beta_\gamma = \delta^\alpha_\gamma$ we have

$$
\begin{aligned}
\frac{\partial}{\partial u^i}(g^\alpha_\beta h^\beta_\gamma) &= \frac{\partial}{\partial u^i}\delta^\alpha_\gamma = 0 \\
\frac{\partial g^\alpha_\beta}{\partial u^i}h^\beta_\gamma + g^\alpha_\beta\frac{\partial h^\beta_\gamma}{\partial u^i} &= 0 \\
\frac{\partial g^\alpha_\beta}{\partial u^i}h^\beta_\gamma &= -g^\alpha_\beta\frac{\partial h^\beta_\gamma}{\partial u^i}
\end{aligned}
$$

and if, for $v = V^\alpha \frac{\partial}{\partial \rho^\alpha} + v^i \frac{\partial}{\partial u^i} = \tilde V^\alpha \frac{\partial}{\partial \tilde\rho^\alpha} + v^i \frac{\partial}{\partial u^i}$ then

$$\begin{aligned}
\tilde V^\beta &= d\tilde\rho^\beta(v) \quad = \quad d(h^\beta_\gamma \rho^\gamma)(v) \\
&= (dh^\beta_\gamma)(v)\rho^\gamma + h^\beta_\gamma\, d\rho^\gamma(v) \\
&= \frac{\partial h^\beta_\gamma}{\partial u^i} du^i(v)\rho^\gamma + h^\beta_\gamma V^\gamma \\
&= \frac{\partial h^\beta_\gamma}{\partial u^i} v^i \rho^\gamma + h^\beta_\gamma V^\gamma
\end{aligned}$$

Note the presence of the first term which may easily be overlooked. It is compensated by ω_G.

Now we can do a standard sort of tensor calculation showing invariance of the coordinate system.

$$\begin{aligned}
D\rho(v) &= \tilde\sigma_\beta\big(\tilde V^\beta + \tilde\Gamma^\beta_{\gamma i} v^i \tilde\rho^\gamma\big) \\
&= \sigma_\alpha g^\alpha_\beta\Big(\frac{\partial h^\beta_\gamma}{\partial u^i}\, v^i\rho^\gamma + h^\beta_\gamma V^\gamma + h^\beta_\delta\, \Gamma^\delta_{\epsilon i}\, g^\epsilon_\gamma v^i\, h^\gamma_\mu \rho^\mu + h^\beta_\mu \frac{\partial g^\mu_\gamma}{\partial u^i}\, h^\gamma_\nu \rho^\nu v^i\Big) \\
&= \sigma_\alpha\Big(g^\alpha_\beta\, \frac{\partial h^\beta_\gamma}{\partial u^i}\, v^i\rho^\gamma + V^\alpha + \Gamma^\alpha_{\mu i}\, v^i\rho^\mu + \frac{\partial g^\alpha_\gamma}{\partial u^i}\, h^\gamma_\nu \rho^\nu v^i\Big) \\
&= \sigma_\alpha\Big(-\frac{\partial g^\alpha_\beta}{\partial u^i} h^\beta_\gamma\, v^i\rho^\gamma + V^\alpha + \Gamma^\alpha_{\mu i}\, v^i\rho^\mu + \frac{\partial g^\alpha_\gamma}{\partial u^i}\, h^\gamma_\nu \rho^\nu v^i\Big) \\
&= \sigma_\alpha\big(V^\alpha + \Gamma^\alpha_{\mu i}\, \rho^\mu v^i\big)
\end{aligned}$$

Since $\tilde\sigma_\beta = \sigma_\alpha g^\alpha_\beta$ we can also express this as

$$\begin{aligned}
g^\alpha_\beta\big(\tilde V^\beta + \tilde\Gamma^\beta_{\gamma i} v^i \tilde\rho^\gamma\big) &= V^\alpha + \Gamma^\alpha_{\delta i}\, \rho^\delta v^i \\
\tilde V^\beta + \tilde\Gamma^\beta_{\gamma i} v^i \tilde\rho^\gamma &= h^\beta_\alpha\big(V^\alpha + \Gamma^\alpha_{\delta i}\, \rho^\delta v^i\big)
\end{aligned}$$

These are the standard tensor coordinate change laws.

There is another way we can look at this; we can look at the coordinates in $\mathbb{R}^n$. To do this we recall

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

We now define an operator Π^α_σ from sections over U to $\mathbb{R}^n$ by

$$\Pi^\alpha_\sigma(v) = V^\alpha + \Gamma^\alpha_{\beta i} v^i \rho^\beta$$

We note that this is the same as $\mathcal{D}_\sigma \rho^\alpha$ defined earlier; the difference is purely psychological. Rewriting in this notation we have

$$\begin{aligned}
\Pi^\beta_{\tilde\sigma}(v) &= \tilde V^\beta + \tilde\Gamma^\beta_{\gamma i} v^i \tilde\rho^\gamma \\
&= h^\beta_\alpha\big(V^\alpha + \Gamma^\alpha_{\delta i}\, v^i \rho^\delta\big) \\
&= h^\beta_\alpha\, \Pi^\alpha_\sigma(v)
\end{aligned}$$

or, in vector form, with $\Pi_\sigma(v) = \left(\Pi^1_\sigma(v), \ldots, \Pi^n_\sigma(v)\right)^\top$,

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

Now for later use we want to write these equations in terms of the $\mathcal{D}$ operator. We have from the above equations

$$\begin{array}{rcl} \mathcal{D}_{\tilde{\sigma}}\tilde{\rho}^\beta(v) & = & h^\beta_\alpha \mathcal{D}_\sigma \rho^\alpha(v) \\ \mathcal{D}_{\sigma g} h^\beta_\alpha \rho^\alpha & = & h^\beta_\alpha \mathcal{D}_\sigma \rho^\alpha \\ \mathcal{D}_{\sigma g} g^{-1} & = & g^{-1}\mathcal{D}_\sigma \\ \mathcal{D}_{\sigma g} & = & g^{-1}\mathcal{D}_\sigma g \quad = \quad \mathrm{Ad}(g^{-1})\mathcal{D}_\sigma \end{array}$$

18. COVARIANT DERIVATIVE OF HIGHER ORDER VECTOR VALUED FORMS

Let $\pi : \mathcal{E} \to M$ be a vector bundle. The covariant derivative D acts on objects in the bundle

$$D\rho \;\; = \;\; \sigma_\alpha \mathcal{D}\rho^\alpha \; = \; \sigma_\alpha(d\rho^\alpha + {\omega_\beta}^\alpha \rho^\beta)$$

The output is a vector valued one form. 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 $\mathcal{A}_i(M)$ consists of 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

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

Vector valued differential forms look locally like

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

We will denote these by

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

Thus the operator D is

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

where the last expression replaces the second as the usual mode of writing for the sake of simplicity.

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:

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

where the last form replaces the second in ordinary calculation. 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.