Connections on principal bundles

Santiago Quintero de los Ríos*

Spring 2019

Contents

1 Connections on principal bundles

1.1 Connections as horizontal distributions

Recall that a vector $v \in T_p P$ is called **vertical** if

$$\mathrm{d}_p \pi(v) = 0.$$

We denote the subspace of vertical vectors by $V_p P \subset T_p P$. By definition, $V_p P$ is nothing more than the kernel of $\mathrm{d}_p \pi$, so we have a short exact sequence

$$0 \longrightarrow V_p P \hookrightarrow T_p P \xrightarrow{\mathrm{d}_p \pi} T_{\pi(p)} M \longrightarrow 0 \ .$$

Since this is a sequence of vector spaces, it splits, and thus we have an isomorphism

$$T_p P \cong V_p P \oplus T_{\pi(p)} M.$$

However, the splitting (and thus the isomorphism) is not canonical: it depends on a choice of a subspace $H_p \subset T_p P$ that is complementary to $V_p P$, and an isomorphism $T_{\pi(p)} M \to H_p$. We call any complementary space to $V_p P$ a **horizontal space** at p, such that:

$$T_p P = V_p P \oplus H_p.$$

1

Once we have chosen a single horizontal subspace $H_p \subset T_p P$ at p, we can find horizontal subspaces for all points in the same fiber of p. This follows since the action of G on P, which we denote $R_g(p) = p \cdot g$, is a fiber-preserving diffeomorphism, and thus $\mathrm{d}_p R_g$ is an isomorphism of tangent spaces that preserves the vertical subspace. This suggests that $\mathrm{d}_p R_g(H_p)$ is a horizontal subspace at $p \cdot g$. Indeed, noting that

$$\mathrm{d}_{p \cdot g} \pi \circ \mathrm{d}_p R_g = \mathrm{d}_p(\pi \circ R_g) = \mathrm{d}_p \pi(v),$$

*Please send corrections, suggestions, etc. to squinterodlr@gmail.com. Latest version on homotopico.com/notes .

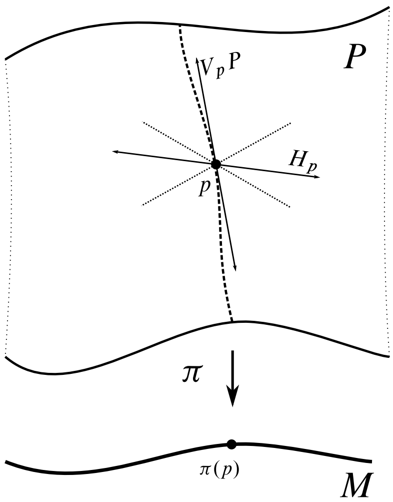

Figure 1: A choice of a horizontal space H_p at T_pP. There are many such choices (in dotted lines).

we see that $\mathrm{d}_p\pi(V_pP) \subseteq V_{p\cdot g}P$. Similarly, if $u \in V_{p\cdot g}P$, we can write

$$u = \mathrm{d}_pR_g(\mathrm{d}_{p\cdot g}R_{g^{-1}}(u)) = \mathrm{d}_pR_g(\tilde{u}),$$

where by the same argument above $\tilde{u} = \mathrm{d}_{p\cdot g}R_{g^{-1}}(u) \in V_pP$ is vertical. Therefore, we obtain that

$$V_{p\cdot g}P = \mathrm{d}_pR_g(V_pP).$$

Furthermore, since $\mathrm{d}_pR_g : T_pP \to T_{p\cdot g}P$ is an isomorphism, we obtain that

$$T_{p\cdot g}P = \mathrm{d}_pR_g(T_pP) = \mathrm{d}_pR_g(V_pP) \oplus \mathrm{d}_pR_g(H_p) = V_{p\cdot g}P \oplus \mathrm{d}_pR_g(H_p),$$

And so we have proved the following:

Lemma 1.1 (Translation of horizontal subspaces). *If $H_p \subset T_pP$ is horizontal at p, then for all $g \in G$, $\mathrm{d}_pR_g(H_p)$ is horizontal at $p \cdot g$.*

So far we have been working at a single point $p \in P$. We can now consider a smooth choice of horizontal spaces above each element of P:

Definition 1.2 ((Principal) Connection). *A **connection** on P is a distribution H on P such that for all $p \in P$, $H_p \subset T_pP$ is a horizontal subspace. We say that a connection H is **principal** if it is compatible with the group action in the sense that for all $g \in G$ and all $p \in P$,*

$$\mathrm{d}_pR_g(H_p) = H_{p\cdot g}. \qquad \triangle$$

The notion of connection is independent of the group action on the total space P, and indeed it applies to general fiber bundles. The condition for a connection to be principal states that our choice of horizontal subspaces along a single fiber is consistent with the "translation" lemma 1.1.

We think of a connection H as a *preferred* way of relating "neighboring" fibers of the bundle. Once we have $p \in P$, we might think that the preferred way of moving to another fiber is along a "direction" (i.e. tangent vector) in the horizontal space H_p. This gives us a little bit of intuition and (sort of) justifies (kind of) the name *connection*. In practice, however, working with distributions might be cumbersome. Fortunately for us, there are other (equivalent) presentations of connections.

1.2 Connections as 1-forms

Let $\mathfrak{g}$ be the Lie algebra of G. Recall that for all $p \in P$, we have the infinitesimal action of $\mathfrak{g}$ on T_pP, $a_p : \mathfrak{g} \to T_pP$ given as

$$a_p(X) := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} p \cdot \exp(tX).$$

Writing $\sigma_p : G \to P$ as $\sigma_p(g) = p \cdot g$, we see that the infinitesimal action is simply the differential of σ_p:

$$a_p(X) = \mathrm{d}_e\sigma_p(X).$$

This infinitesimal action induces, for each $X \in \mathfrak{g}$, a vector field $X^\sharp$ called the **fundamental vector field** associated to X given by

$$X^\sharp_p := a_p(X).$$

We have that σ_p is a diffeomorphism onto the fiber containing p, and thus $a_p = \mathrm{d}_e\sigma_p$ induces a linear isomorphism $\mathfrak{g} \overset{a_p}{\cong} V_pP$.

Suppose that we have a principal connection H on P. Then in particular, we have a subspace $H_p \subset T_pP$ such that $T_pP = V_pP \oplus H_p$, and so we can construct a map $\omega_p : T_pP \to \mathfrak{g}$ as

$$\omega_p(v^V + v^H) = a_p^{-1}(v^V),$$

where $v^V \in V_pP$ and $v^H \in H_p$. By construction, we have that

$$\omega_p(a_p(X)) = X$$

for all $X \in \mathfrak{g}$. We can also see how ω_p compares to $\omega_{p\cdot g}$, since we know that our horizontal distribution behaves nicely along the fibers of the action.

For this, first note that for all $g \in G$,

$$\begin{aligned}
\mathrm{d}_pR_g(a_p(X)) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} R_g(p\cdot\exp(tX)) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} p\cdot\exp(tX)g \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (p\cdot g)\cdot(g^{-1}\exp(tX)g).
\end{aligned}$$

Now we ask ourselves, do we know what the tangent vector of $g^{-1}\exp(tX)g$ is? Yes, yes we do:

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g^{-1}\exp(tX)g = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{Conj}_{g^{-1}}(\exp(tX)) = \mathrm{Ad}_{g^{-1}}(X),$$

where we have written[1] $\mathrm{Conj}_g(h) = ghg^{-1}$, and $\mathrm{Ad}_g = \mathrm{d}_e\,\mathrm{Conj}_g$. Then we have

$$\mathrm{d}_pR_g(a_p(X)) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (p\cdot g)\cdot(g^{-1}\exp(tX)g) = a_{p\cdot g}(\mathrm{Ad}_{g^{-1}}(X)).$$

With this, we can see that for $v \in T_pP$, which we write as $v = v^V + v^H$ with $v^V = a_p(X)$ for some $X \in \mathfrak{g}$:

$$(R_g^*\omega)_p(v^V + v^H) = \omega_{p\cdot g}(\mathrm{d}_pR_g(v^V) + \mathrm{d}_pR_g(v^H)) = \omega_{p\cdot g}(\mathrm{d}_gR_g(a_p(X))) = \mathrm{Ad}_{g^{-1}}(X) = (\mathrm{Ad}_{g^{-1}}\circ\omega_p)(v),$$

and so we conclude that

$$(R_g^*\omega) = \mathrm{Ad}_{g^{-1}}\circ\omega.$$

Then we have proved, modulo the small detail of smoothness[2], the following:

Proposition 1.3 (1-form induced by principal connection). *Let H be a principal connection on $G \hookrightarrow P \xrightarrow{\pi} M$. Then there exists a (unique) $\mathfrak{g}$-valued 1-form $\omega \in \Omega^1(P,\mathfrak{g})$, such that for all $p \in P$, $g \in G$ and $X \in \mathfrak{g}$:*

1. $\omega_p(a_p(X)) = X$,
2. $R_g^*\omega = \mathrm{Ad}_{g^{-1}}\circ\omega$, *and*
3. $\ker(\omega_p) = H_p$.

We call any $\mathfrak{g}$-valued 1-form satisfying these properties a **connection 1-form**:

Definition 1.4 (Connection 1-form). *A connection 1-form on P is a $\mathfrak{g}$-valued 1-form $\omega \in \Omega^1(P,\mathfrak{g})$ such that for all $p \in P$, $g \in G$ and $X \in \mathfrak{g}$:*

1. $\omega_p(a_p(X)) = X$, *and*

[1] `https://xkcd.com/927/`
[2] We can handwave it away by saying that it follows from the smoothness of the distribution H.

2. $R_g^*\omega = \mathrm{Ad}_{g^{-1}} \circ \omega$. △

The converse to proposition 1.3 is also true:

Proposition 1.5. *Principal connection induced by connection 1-form Let $\omega \in \Omega^1(P, \mathfrak{g})$ be a connection 1-form. Then the distribution H defined pointwise as*

$$H_p = \ker(\omega_p) \subset T_p P$$

is a principal connection on P.

Proof.— First, let's see that indeed $H_p = \ker(\omega_p)$ is horizontal. If $v \in \ker(\omega_p) \cap V_p P$, then $v = a_p(X)$ for some $X \in \mathfrak{g}$, so that

$$0 = \omega_p(v) = \omega_p(a_p(X)) = X,$$

and thus $v = 0$. Therefore $\ker(\omega_p) \cap V_p P = \{0\}$. Now for an arbitrary $v \in T_p P$, set

$$v^V = a_p(\omega_p(v)).$$

Then we have that $\mathrm{d}_p\pi(v^V) = 0$, since it is in the image of a_p, and thus $v^V \in V_p P$. Finally, setting $v^H = v - v^V$, we have

$$\omega_p(v^H) = \omega_p(v) - \omega_p(a_p(\omega_p(v))) = \omega_p(v) - \omega_p(v) = 0,$$

and so $v^H \in \ker\omega_p = H_p$. We have then shown that $v = v^V + v^H$, with $v^V \in V_p P$ and $v^H \in H_p$, and so

$$T_p P = V_p P \oplus H_p.$$

Thus H_p is a horizontal subspace. Now to see that H is principal, note that

$$\omega_{p\cdot g}(\mathrm{d}_p R_g(v)) = \mathrm{Ad}_{g^{-1}}(\omega_p(v)).$$

Since both $\mathrm{d}_p R_g$ and $\mathrm{Ad}_{g^{-1}}$ are isomorphisms, we have that $v \in \ker\omega_p$ if and only if $\mathrm{d}_p R_g(v) \in \ker\omega_{p\cdot g}$, and thus

$$\mathrm{d}_p R_g(H_p) = H_{p\cdot g}.$$

Finally, smoothness follows from the fact that ω is a smooth form. ■

From now on, if ω is a connection 1-form, we will simply call it a connection. In physics lingo, connections are often called *gauge fields* or *gauge potentials*.

Example 1.6 (Maurer-Cartan connection). Let G be a Lie group, which we interpret as a principal G-bundle over a one-point space $G \hookrightarrow G \xrightarrow{\pi} \{\star\}$. For each $g \in G$, we have a way to map $T_g G$ to $\mathfrak{g} = T_e G$, simply by pushing vectors via one of the multiplications; for instance

$$\mathrm{d}_g L_{g^{-1}} : T_g G \to \mathfrak{g} = T_e G.$$

We then define the **Maurer-Cartan** form of G, denoted by $\Theta \in \Omega^1(G, \mathfrak{g})$, as

$$\Theta_g = \mathrm{d}_g L_{g^{-1}}.$$

The heading of the example spoiled the surprise: Θ is a connection on G. Indeed, for $X \in \mathfrak{g} = T_e G$, we have that

$$a_g(X) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g\exp(tX) = \mathrm{d}_e L_g(X),$$

so that

$$\Theta_g(a_g(X)) = \mathrm{d}_g(L_{g^{-1}})(\mathrm{d}_e L_g(X)) = \mathrm{d}_g(L_{g^{-1}} \circ L_g)(X) = X.$$

Now for any $h \in G$, we have

$$(R_g^*\Theta)_h(X) = \Theta_{hg}(\mathrm{d}_h R_g(X)) = \mathrm{d}_{hg} L_{g^{-1}h^{-1}} \mathrm{d}_h R_g(X) = \mathrm{d}_h(L_{g^{-1}h^{-1}} \circ R_g)(X).$$

But then, we see that

$$(L_{g^{-1}h^{-1}} \circ R_g)(x) = g^{-1}h^{-1}xg = (\mathrm{Conj}_{g^{-1}} \circ L_{h^{-1}})(x),$$

such that the differential at h is

$$\mathrm{d}_h(L_{g^{-1}h^{-1}} \circ R_g) = \mathrm{d}_h(\mathrm{Conj}_{g^{-1}} \circ L_{h^{-1}}) = \mathrm{d}_e \mathrm{Conj}_{g^{-1}} \mathrm{d}_h L_{h^{-1}} = \mathrm{Ad}_{g^{-1}} \circ \Theta_h,$$

and so, indeed

$$(R_g^*\Theta) = \mathrm{Ad}_{g^{-1}} \circ \Theta.$$ ♣

With the Maurer-Cartan form, we can construct connections on any principal bundle.

Example 1.7 (Trivial connection on a trivial bundle). Let $P = M \times G$ be a trivial bundle, and $\mathrm{pr}_2 : M \times G \to G$ be the projection onto G. If Θ is the Maurer-Cartan form of G, then $\mathrm{pr}_2^*\Theta$ is a connection on $M \times G$, and its horizontal distribution is precisely given by $H_{(x,g)} := T_x M \oplus 0 \subset T_{(x,g)}(M \times G)$. ♣

Example 1.8 (Flat connection on a bundle). Let $G \hookrightarrow P \xrightarrow{\pi} M$ be a principal G-bundle, with a trivializing cover $\{(U_j, \Psi_j)\}_{j \in J}$. We write each $\Psi_j : \pi^{-1}(U_j) \to U_j \times G$ as

$$\Psi_j(p) = (\pi(p), \psi(p)),$$

with $\psi : \pi^{-1}(U_j) \to G$. Then $\psi_j^*\Theta$ is a connection on the trivial bundle $\pi^{-1}(U_j) \cong U_j \times G$, where Θ is the Maurer-Cartan form of G. Now let $\{f_j\}_{j \in J}$ be a partition of unity subordinate to $\{U_j\}_{j \in J}$. Then the form defined as

$$\omega = \sum_j (f_j \circ \pi)\psi_j^*\Theta$$

is a connection on P. ♣

1.3 Local expressions, or, why physicists did nothing wrong

Consider a trivializing cover $\{(U_j, \Psi_j)\}_{j \in J}$ of the bundle $\pi : P \to M$, where we write each $\Psi_i : \pi^{-1}(U_i) \to U_i \times G$ as

$$\Psi_i(p) = (\pi(p), \psi_i(p)),$$

with $\psi_i : U_i \to G$. We know that each trivialization Ψ_i has an associated section $s_i : U_i \to P$, given by

$$s_i(x) = \Psi_i^{-1}(x, e)$$

for all $x \in U_i$. These sections are called **local gauges** in the physics literature.

Note that for all $x \in U_i$ and $p \in \pi^{-1}(x)$,

$$\Psi_i(s_i(x) \cdot \psi_i(p)) = (x, \psi_i(s_i(x))\psi_i(p)) = (x, \psi_i(p)) = \Psi_i(p),$$

and therefore we have that

$$p = s_i(x) \cdot \psi_i(p).$$

Now if $x \in U_{ij} = U_i \cap U_j$, for all elements $p \in \pi^{-1}(x)$, we obtain for both sections

$$s_i(x) \cdot \psi_i(p) = p = s_j(x) \cdot \psi_j(p),$$

and thus

$$s_j(x) = s_i(x) \cdot \psi_i(p)\psi_j(p)^{-1}.$$

But now, since the trivializations are G-equivariant, $\psi_i(p \cdot g) = \psi_i(p)g$, the product $\psi_i(p)\psi_j(p)^{-1}$ is G-invariant, and is precisely the transition function $g_{ij} : U_{ij} \to G$:

$$g_{ij}(x) := \psi_i(p)\psi_j(p)^{-1}.$$

We then conclude:

$$s_j(x) = s_i(x) \cdot g_{ij}(x).$$

See fig. 2.

Now let $\omega \in \Omega^1(P, \mathfrak{g})$ be a connection. For each U_i, the pullback of ω by s_i is again a $\mathfrak{g}$-valued 1-form on U_i. We denote it by

$$\omega_i := s_i^*\omega$$

and call it the **local gauge potential** (in the gauge s_i). How do different local gauges relate to one another?

Proposition 1.9 (Transformation of local potentials). *Let ω be a connection on $G \hookrightarrow P \xrightarrow{\pi} M$, and $\{U_i\}_{i \in J}$ a trivializing cover with induced sections $s_i : U_i \to P$, and transition maps $g_{ij} : U_i \cap U_j \to G$. Let $\omega_i = s_i^*\omega$ be the local gauge potentials. Then for all $x \in U_{ij} = U_i \cap U_j$,*

$$(\omega_j)_x = \mathrm{Ad}_{g_{ij}(x)^{-1}} \circ (\omega_i)_x + (g_{ij}^*\Theta)_x, \tag{1}$$

where Θ is the Maurer-Cartan form of example 1.6.

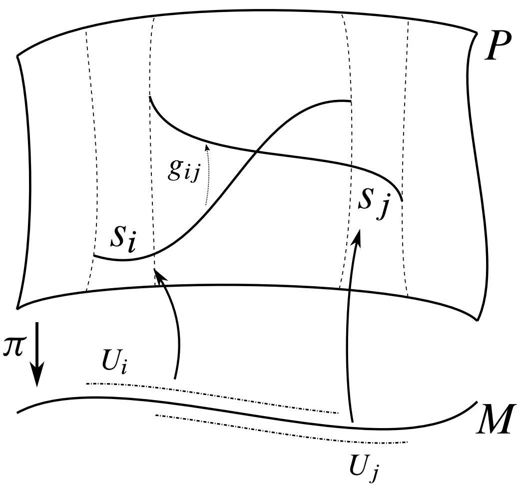

Figure 2: The transition functions g_{ij} relate the sections s_i, s_j induced by the trivializations.

Proof. — Let's try to brute-force it first, and see what else we need. we have that

$$(\omega_j)_x = (s_j^*\omega)_x = \omega_{s_j(x)} \circ \mathrm{d}_x s_j,$$

so we need to find the expression for $\mathrm{d}_x s_j$, preferably in terms of s_i. To do so, let $\sigma : P \times G \to P$ be the action, i.e. $\sigma(p, g) = p \cdot g$. Then for all $x \in U_{ij}$ we can write $s_j(x)$ as

$$s_j(x) = s_i(x) \cdot g_{ij}(x) = \sigma(s_i(x), g_{ij}(x)) = (\sigma \circ (s_i, g_{ij}))(x),$$

where we have $(s_j, g_{ij}) : U \to P \times M$ is defined in the natural way. This tells us that

$$\mathrm{d}_x s_j = \mathrm{d}_x(\sigma \circ (s_i, g_{ij})) = \mathrm{d}_{(s_j(x), g_{ij}(x))}\sigma \circ \mathrm{d}_x(s_j, g_{ij}) = \mathrm{d}_{(s_j(x), g_{ij}(x))}\sigma \circ (\mathrm{d}_x s_j, \mathrm{d}_x g_{ij}).$$

Now we need to find the expression for $\mathrm{d}_{(p,g)}\sigma$. We proceed carefully, in parts, noting that $T_{(p,g)}(P \times G) \cong T_p P \oplus T_g G$. Let $u \in T_p P$, and γ an integral curve of u. Then we have that

$$\mathrm{d}_{(p,g)}\sigma(u, 0) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \sigma(\gamma(t), g) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \gamma(t) \cdot g = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} R_g(\gamma(t)) = \mathrm{d}_p R_g(u).$$

On the other hand, let $\xi \in T_g G$. Then we have that $\Theta_g(\xi) := X \in \mathfrak{g} = T_e G$ is the (unique) element of the Lie algebra that satisfies

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g \exp(tX) = \mathrm{d}_e L_g(X) = \xi,$$

so that $t \mapsto g \exp\big(t\Theta_g(\xi)\big)$ is an integral curve of ξ. Therefore

$$\begin{aligned}
\mathrm{d}_{(p,g)}\sigma(0, \xi) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \sigma(p, g \exp\big(t\Theta_g(\xi)\big)) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} p \cdot g \exp\big(t\Theta_g(\xi)\big) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (p \cdot g) \cdot \exp\big(t\Theta_g(\xi)\big) \\
&= a_{p \cdot g}(\Theta_g(\xi)).
\end{aligned}$$

We put these two together, and obtain

$$\mathrm{d}_{(p,g)}\sigma(u, \xi) = \mathrm{d}_p R_g(u) + a_{p \cdot g}(\Theta_g(\xi)).$$

Substituting in $\mathrm{d}_x s_j$, and evaluating at some $v \in T_x U_{ij}$,

$$\begin{aligned}
\mathrm{d}_x s_j(v) = \mathrm{d}_{(s_i(x), g_{ij}(x))}\sigma(\mathrm{d}_x s_j(v), \mathrm{d}_x g_{ij}(v)) &= \mathrm{d}_{s_i(x)} R_{g_{ij}(x)}(\mathrm{d}_x s_i(v)) + a_{s_i(x) \cdot g_{ij}(x)}(\Theta_{g_{ij}(x)}(\mathrm{d}_x g_{ij}(v))). \\
&= \mathrm{d}_{s_i(x)} R_{g_{ij}(x)}(\mathrm{d}_x s_i(v)) + a_{s_j(x)}((g_{ij}^*\Theta)_x(v)).
\end{aligned}$$

Now we evaluate $\omega_{s_j(x)}$ on $\mathrm{d}_x s_j(v)$. By definition, we have

$$\omega_{s_j(x)}(a_{s_i(x)}((g_{ij}^*\Theta)_x(v))) = (g_{ij}^*\Theta)_x(v).$$

We have to do a little bit more work for the other term. We will simply write s_j, g_{ij}, s_j for $s_j(x)$, etc., to avoid the clutter. Then we have

$$\begin{aligned}
\omega_{s_j}(\mathrm{d}_x s_j(u)) &= \omega_{s_j}(\mathrm{d}_{s_i} R_{g_{ij}}(\mathrm{d}_x s_i(u))) \\
&= \omega_{s_i g_{ij}}(\mathrm{d}_{s_i} R_{g_{ij}}(\mathrm{d}_x s_i(u))) \\
&= (R_{g_{ij}}^*\omega)_{s_i}(\mathrm{d}_x s_i(u)) \\
&= \mathrm{Ad}_{g_{ij}^{-1}}(\omega_{s_i}(\mathrm{d}_x s_i(u))) \\
&= \mathrm{Ad}_{g_{ij}^{-1}}((s_i^*\omega)_x(u)) \\
&= (\mathrm{Ad}_{g_{ij}(x)^{-1}} \circ (\omega_i)_x)(u).
\end{aligned}$$

Placing these two last results together, we obtain the result. ■

This proposition, in physics, is often called *gauge transformation* of a potential. In physics we mostly with the local potentials, not with the global connection in the total space P, and we define a gauge potential as *some* object that under a certain set of (local) transformations, transforms as in eq. (1). Indeed, the following result tells us that this information is sufficient to reconstruct the global object. The proof is a bit tedious and not particularly enlightening (we did a lot of the work in previous proposition).

Proposition 1.10 (Physicists did nothing wrong). *Let $G \hookrightarrow P \xrightarrow{\pi} M$ be a principal G-bundle, and $\{(U_i, \Psi_i)\}_{i\in J}$ a trivializing cover with induced sections $s_i : U_i \to P$. Suppose that for each U_i, there is a $\mathfrak{g}$-valued 1-form $\omega_i \in \Omega^1(U_i, \mathfrak{g})$, such that for all $x \in U_i \cap U_j$,*

$$(\omega_j)_x = \mathrm{Ad}_{g_{ij}(x)^{-1}} \circ (\omega_i)_x + g_{ij}^*\Theta_x.$$

Then there exists a unique connection $\omega \in \Omega^1(P, \mathfrak{g})$ such that for all $i \in J$,

$$s_i^*\omega = \omega_i.$$

1.4 Horizontal lifts

Once we have a connection, we now have a preferred way of *lifting* vectors from TM to TP. Recall that a vector $Y \in T_pP$ is a **lift** of $X \in T_{\pi(p)}M$ if $\mathrm{d}_p\pi(Y) = X$. In absence of a connection, there are many different choices of lifts of a vector, and any two choices differ by a vertical vector. That is, if Y, Y' are lifts of X, then $Y - Y'$ is vertical. Once we have a connection, we can define the **horizontal lift** (with respect to a connection H) of $X \in T_xM$ as the horizontal component of *any* lift of X. This definition is, of course, independent of the choice of lift, since any two differ by a vertical vector, whose horizontal component vanishes. Denoting the horizontal component of a vector by Y^H, we have then

$$Y^H = (Y' + (Y - Y'))^H = (Y')^H.$$

Similarly, we can lift vector fields by lifting them in a pointwise fashion.

Definition 1.11 (Horizontal lift of vector fields). *Let $X \in \mathfrak{X}(M)$ be a vector field. We define the **horizontal lift** of X as a vector field $Y \in \mathfrak{X}(P)$, where Y_p is the horizontal lift of $X_{\pi(p)}$.* △

If H is a principal connection, then the horizontal lift Y of a vector field X is G-invariant, since $\mathrm{d}_pR_g(Y_p)$ is a horizontal vector that projects to $X_{\pi(p)}$. Therefore we have that

$$R_{g*}Y = Y.$$

We also expect a horizontal lift to commute with (some) vertical fields, since, in a sketchy intuitive sense, we define these two directions as independent. Actually, this is true of any G-invariant field.

Lemma 1.12 (G-invariant fields commute with fundamental vector fields). . *Let $X^\sharp \in \mathfrak{X}(P)$ be the fundamental vector field associated to $X \in \mathfrak{g}$, and let $Y \in \mathfrak{X}(P)$ be a G-invariant field, i.e. $R_{g*}Y = Y$. Then $[X, Y] = 0$.*

Proof.— Let Φ_t be the flow of $X^\sharp$. It is straightforward to check that

$$\Phi_t(p) = p \cdot \exp(tX) = R_{g_t}(p),$$

where we denote $g_t = \exp(tX)$. Then

$$[X, Y]_p = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{d}_{\Phi_t(p)}\Phi_{-t}(Y_{\Phi_t(p)}) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{d}_{p\cdot g_t}R_{g_t^{-1}}(Y_{p\cdot g_t}) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} Y_p = 0.$$ ■

2 Curvature

2.1 The curvature 2-form and structure equation

Let $G \hookrightarrow P \overset{\pi}{\to} M$ be a principal G-bundle, and $\mathfrak{g}$ be the Lie algebra of G. For any $\mathfrak{g}$-valued k-form $\omega \in \Omega^k(P, \mathfrak{g})$, we define $\mathrm{d}\omega \in \Omega^{k+1}(P, \mathfrak{g})$ as follows: choose a basis $\{e_1, \ldots, e_m\}$ of $\mathfrak{g}$. Then we can write

$$\omega = \sum_{a=1}^{m} \omega^a e_a,$$

where each $\omega^a \in \Omega^k(P)$. Then we define

$$\mathrm{d}\omega := \sum_{a=1}^{m} \mathrm{d}\omega^a e_a.$$

This definition is independent of the choice of basis of $\mathfrak{g}$, as can be readily checked.

In order to define curvature, we also need another definition.

Definition 2.1 (Bracket of valued forms). *Let $\alpha \in \Omega^k(P, \mathfrak{g})$ and $\beta \in \Omega^l(P, \mathfrak{g})$. We define a $(k+l)$-form $[\alpha, \beta] \in \Omega^{k+l}(P, \mathfrak{g})$ in terms of a basis $\{e_1, \ldots, e_m\}$ of $\mathfrak{g}$ as*

$$[\alpha, \beta] = \sum_{a,b} \alpha^a \wedge \beta^b [e_a, e_b].$$

This definition is independent of the choice of basis (and in some references it is written as $\alpha \wedge \beta$). △

In the case where $\alpha, \beta \in \Omega^1(P, \mathfrak{g})$, the definition becomes

$$\begin{aligned}
[\alpha, \beta](X, Y) &= \sum_{a,b} (\alpha^a \wedge \beta^b)(X, Y)[e_a, e_b] \\
&= \sum_{a,b} (\alpha^a(X)\beta^b(Y) - \alpha^a(Y)\beta^b(X))[e_a, e_b] \\
&= \sum_{a,b} [\alpha^a(X)e_a, \beta^b(Y)e_b] - [\alpha^a(Y)e_a, \beta^b(X)e_b] \\
&= [\alpha(X), \beta(Y)] - [\alpha(Y), \beta(X)].
\end{aligned}$$

In general, we have a way to evaluate the bracket of forms:

Lemma 2.2 (Evaluation of bracket). *Let $\alpha \in \Omega^i(P, \mathfrak{g})$ and $\beta \in \Omega^j(P, \mathfrak{g})$. Then for vectors $X_1, \ldots, X_{i+j}$:*

$$[\alpha, \beta](X_1, \ldots, X_{i+j}) = \frac{1}{i!j!} \sum_{\sigma \in \mathfrak{S}_k} \operatorname{sgn}(\sigma)[\alpha(X_{\sigma(1)}, \ldots, X_{\sigma(i)}), \beta(X_{\sigma(i+1)}, \ldots, X_{\sigma(i+j)})].$$

Proof. — The proof is a straightforward evaluation and application of the definition of the wedge product. ■

Now we define the curvature 2-form of a connection.

Definition 2.3 (Curvature 2-form). *Let ω be a connection on $G \hookrightarrow P \overset{\pi}{\to} M$. The **curvature** of ω is a $\mathfrak{g}$-valued 2-form $\Omega \in \Omega^2(P, \mathfrak{g})$ defined as*

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

*Equation (2) is called the **Cartan structure equation**.* △

Now let's do an example that comes with a bit of motivation.

Example 2.4 (Curvature of the Maurer-Cartan form). Recall that for a Lie group G (which we see as a G-bundle over a one-point space), we have a canonical connection Θ on G, called the Maurer-Cartan form (see example 1.6), given pointwise as

$$\Theta_g = \mathrm{d}_g(L_{g^{-1}}).$$

The Maurer-Cartan form is also left-invariant,

$$(L_g^*\Theta)_h = \Theta_{gh} \circ \mathrm{d}_h L_g = \mathrm{d}_{gh} L_{(gh)^{-1}} \mathrm{d}_h L_g = \mathrm{d}_h(L_{(gh)^{-1}} \circ L_g) = \mathrm{d}_h L_{h^{-1}} = \Theta_h,$$

and so it is uniquely defined by its value at the identity $e \in G$. Now fix a basis $\{e_1, \dots, e_m\}$ of $\mathfrak{g}$, so that the Maurer-Cartan form can be written as

$$\Theta = \sum_a \Theta^a e_a,$$

where each component $\Theta^a \in \Omega^1(G)$ is a usual 1-form. If we write ξ_a as the left-invariant field generated by e_a;

$$\xi_a(g) := \mathrm{d}_e L_g(e_a),$$

then at each point $g \in G$, the set $\{\xi_1(g), \dots, \xi_m(g)\}$ are a frame for $T_g G$, and furthermore, we have that

$$\Theta_g(\xi_a(g)) = e_a.$$

But on the other hand, we have

$$\Theta_g(\xi_a(g)) = \sum_b \Theta_g^b(\xi_a(g)) e_a,$$

which implies that for all g, $\Theta_g^b(\xi_a(g)) = \delta_a^b$, and so $\{\Theta^1, \dots, \Theta^m\}$ forms a coframe of $T_g^* G$ that is dual to $\{\xi_1, \dots, \xi_m\}$. Now we have that

$$\mathrm{d}\Theta^a(\xi_b, \xi_c) = \xi_b(\Theta^a(\xi_c)) - \xi_c(\Theta^a(\xi_b)) - \Theta^a([\xi_b, \xi_c]) = -\Theta_e^a([e_b, e_c]) = -[e_b, e_c]^a = -C_{bc}^a,$$

where $[e_b, e_c]^a$ is the a-th component of $[e_b, e_c]$, which is precisely the definition of the structure coefficients C_{bc}^a. Now since the ξ_a vectors form a frame of TG, whose dual coframe is precisely the Θ^a forms, this tells us that

$$\mathrm{d}\Theta^a = -\frac{1}{2} \sum_{b,c} C_{bc}^a \Theta^b \wedge \Theta^c,$$

and thus,

$$\mathrm{d}\Theta = \sum_a \mathrm{d}\Theta^a e_a = -\frac{1}{2} \sum_{a,b,c} C_{bc}^a \Theta^b \wedge \Theta^c e_a = -\frac{1}{2} \sum_{b,c} \Theta^b \wedge \Theta^c [e_b, e_c] = -\frac{1}{2}[\Theta, \Theta].$$

Therefore, conclude that

$$\Omega = \mathrm{d}\Theta + \frac{1}{2}[\Theta, \Theta] = 0. \qquad \clubsuit$$

From this example, we can see that the curvature of the flat connection of example 1.8 has vanishing curvature as well.

Example 2.5 (Flat connection induced by Maurer-Cartan has vanishing curvature). Let ω be the flat connection of $G \hookrightarrow P \xrightarrow{\pi} M$, as in example 1.8. Then on each $\pi^{-1}(U_j)$, writing $\omega_j = \omega|_{\pi^{-1}(U_j)}$ we have

$$\Omega|_{\pi^{-1}(U_j)} = \mathrm{d}\omega_j + \frac{1}{2}[\omega_j, \omega_j] = \mathrm{d}(\psi_j^* \Theta) + \frac{1}{2}[\psi_j^* \Theta, \psi_j^* \Theta] = \psi_j^*(\mathrm{d}\Theta + \frac{1}{2}[\Theta, \Theta]) = 0. \qquad \clubsuit$$

We now see one of the most (if not the most) important properties of the curvature 2-form:

Proposition 2.6 (Curvature is basic). *Let ω be a connection on $G \hookrightarrow P \xrightarrow{\pi} M$ and Ω its curvature. Then $\Omega \in \Omega_{bas}^2(P, \mathfrak{g})$, that is,*

1. *If X is a vertical field, then $\iota_X \Omega = 0$, i.e. Ω is horizontal; and*
2. *For all $g \in G$, $R_g^* \Omega = \mathrm{Ad}_{g^{-1}} \circ \Omega$, i.e. Ω is pseudotensorial[3] of type* Ad.

Proof. — First, let's see that Ω is pseudotensorial of type Ad:

$$R_g^* \Omega = R_g^* \mathrm{d}\omega + \frac{1}{2} R_g^*[\omega, \omega] = \mathrm{d}R_g^* \omega + \frac{1}{2}[R_g^* \omega, R_g^* \omega] = \mathrm{d}(\mathrm{Ad}_{g^{-1}} \circ \omega) + \frac{1}{2}[\mathrm{Ad}_{g^{-1}} \circ \omega, \mathrm{Ad}_{g^{-1}} \circ \omega].$$

The occurrences of $\mathrm{Ad}_{g^{-1}}$ in this previous expression may seem like there's some care required with d and the commutator, but by definition, $\mathrm{Ad}_{g^{-1}}$ acts on the element of $\mathfrak{g}$ that ω outputs. We can see this more clearly when we choose a basis $\{e_1, \dots, e_m\}$ of $\mathfrak{g}$ and write $\omega = \sum_a \omega^a e_a$. When we write $\mathrm{Ad}_{g^{-1}} \circ \omega$, this actually stands for

$$\mathrm{Ad}_{g^{-1}} \circ \omega = \sum_a \omega^a \mathrm{Ad}_{g^{-1}}(e_a),$$

[3]this is sometimes called G-invariance, or G-equivariance, but let's avoid that discussion.

so that the terms in the previous expression are

$$\mathrm{d}\left(\mathrm{Ad}_{g^{-1}} \circ \sum_a \omega^a e_a\right) = \sum_a \mathrm{d}\omega^a \mathrm{Ad}_{g^{-1}}(e_a) = \mathrm{Ad}_{g^{-1}} \circ \mathrm{d}\omega.$$

Now we consider the case of the bracket. Since $\mathrm{Ad}_g = \mathrm{d}_e C_g$ is the differential of a diffeomorphism, it is a pushworward evaluated at e and thus it distributes into the Lie bracket of vector fields

$$\begin{aligned}
\mathrm{Ad}_g[X_e, Y_e] &= (\mathrm{Conj}_{g*}[X, Y])_e \\
&= [\mathrm{Conj}_{g*} X, \mathrm{Conj}_{g*} Y]_e \\
&= [\mathrm{Ad}_g(X_e), \mathrm{Ad}_g(Y_e)].
\end{aligned}$$

Then we have

$$(R_g^* \Omega) = \mathrm{Ad}_{g^{-1}}\left(\mathrm{d}\omega + \frac{1}{2}[\omega, \omega]\right) = \mathrm{Ad}_{g^{-1}} \circ \Omega.$$

We now need to show that Ω is horizontal. Since we have a connection, we can decompose any vector $v \in T_p P$ in a vertical and horizontal part, $v = v^V + v^H$. Then the action on Ω on a pair $u, v \in T_p P$ is

$$\Omega_p(u, v) = \Omega_p(u^V + u^H, v^V + v^H) = \Omega_p(u^V, v^V) + \Omega_p(u^V, v^H) + \Omega_p(u^H, v^V) + \Omega_p(u^H, v^H),$$

and thus, it suffices to consider two cases: when both u and v are vertical, or when u is vertical and v is horizontal.

Let's begin with with the case where both u and v are vertical, so that $u = a_p(X)$ and $v = a_p(Y)$ for some $X, Y \in \mathfrak{g}$ (namely $X = \omega_p(u)$ and $Y = \omega_p(v)$). If we write $X^\sharp, Y^\sharp$ for the fundamental vector fields associated to X, Y, given by $X_p^\sharp = a_p(X) = \sigma_{p*}(X)$ (and same for Y), we have then that

$$\begin{aligned}
\Omega_p(u, v) &= \mathrm{d}\omega_p(u, v) + \frac{1}{2}[\omega, \omega](u, v) \\
&= u(\omega(Y^\sharp)) - v(\omega(X^\sharp)) - \omega([u, v]) + [\omega(u), \omega(v)].
\end{aligned}$$

But $\omega(X^\sharp) = X$ and $\omega(Y^\sharp) = Y$ are constant, so

$$\Omega_p(u, v) = -\omega([u, v]) + [X, Y].$$

Finally, we see that

$$[u, v] = [X^\sharp, Y^\sharp]_p = [\sigma_{p*}X, \sigma_{p*}Y]_p = \sigma_{p*}([X, Y]_e) = a_p([X, Y]),$$

so $\omega([u, v]) = [X, Y]$, and thus

$$\Omega_p(u, v) = 0.$$

Now let's consider the case where u is vertical and v is horizontal. Again, let $X = \omega_p(u) \in \mathfrak{g}$, and $X^\sharp$ be the fundamental vector field associated to X, so that $X_p^\sharp = u$; and let ν be a horizontal field such that $\nu_p = v$. We then have

$$\begin{aligned}
\Omega_p(u, v) &= \mathrm{d}\omega_p(u, v) + \frac{1}{2}[\omega, \omega](u, v) \\
&= u(\omega(\nu)) - v(\omega(X^\sharp)) - \omega([u, v]) + [\omega(u), \omega(v)] \\
&= -\omega([u, v]).
\end{aligned}$$

Now it suffices to show that $[u, v]$ is horizontal if v is horizontal and u is vertical. First, we have that the flow of the fundamental vector field $X^\sharp$ is given by

$$\Phi_t(p) = p \cdot \exp(tX),$$

as can be readily checked. Then

$$[u, v] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\Phi_{-t*}(\nu))_p = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{d}_{\Phi_t(p)} \Phi_{-t}(\nu_{\Phi_t(p)})$$

If we write $g_t = \exp(tX)$, then it is clear that $\Phi_t(p) = R_{g_t}(p)$, so

$$[u, v] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{d}_{\Phi_t(p)} \Phi_{-t}(\nu_{\Phi_t(p)}) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{d}_{p \cdot g_t} R_{g_t^{-1}}(\nu_{p \cdot g_t}).$$

However, we know that $(R_{g*})(v)$ is horizontal for all g if v is horizontal, and thus we obtain that

$$\mathrm{d}_{p \cdot g_t} R_{g_t^{-1}}(\nu_{p \cdot g_t}) \in H_p \qquad \text{for all } t,$$

and so $[u, v]$ is horizontal as well. Therefore $\omega([u, v]) = 0$, and our result is proved. ■

Since Ω is horizontal, its values are uniquely determined by the horizontal components of the vectors that it is evaluated at. The following corollary is often given as the definition of the curvature form:

Corollary 2.7. *Let ω be a connection and Ω its curvature. Then for all $u, v \in TP$:*

$$\Omega(u, v) = \mathrm{d}\omega(u^H, v^H),$$

where u^H, v^H are the horizontal components of u, v, determined by ω.

2.2 The exterior covariant derivative

From corollary 2.7, we see that the curvature Ω can be defined as the horizontal component of $\mathrm{d}\omega$. We can extend this notion, and define the **exterior covariant derivative** $\mathrm{d}^\omega : \Omega^k(P, \mathfrak{g}) \to \Omega^{k+1}(P, \mathfrak{g})$ as the horizontal component of the usual de Rham differential:

$$\mathrm{d}^\omega\alpha(X_1, \dots, X_{k+1}) := \mathrm{d}\alpha(X_1^H, \dots, X_{k+1}^H).$$

With this definition, we can simply write

$$\Omega = \mathrm{d}^\omega\omega.$$

Clearly, by definition, $\mathrm{d}^\omega\alpha$ is horizontal for any form $\alpha \in \Omega^k(P, \mathfrak{g})$. We also see that $\mathrm{d}^\omega\alpha$ is pseudotensorial of type Ad if α also is. The idea is that R_g preserves horizontality and the pullback commutes with d, so in general pulling back by R_g should behave reasonable well. Indeed, let $\alpha \in \Omega^k(P, \mathfrak{g})$ be pseudotensorial of type Ad. Then

$$\begin{aligned}
(R_g^*\mathrm{d}^\omega\alpha)_p(X_1, \dots, X_{k+1}) &= (\mathrm{d}^\omega\alpha)_{p\cdot g}(R_{g*}X_1, \dots, R_{g*}X_{k+1}) \\
&= \mathrm{d}\alpha_{p\cdot g}((R_{g*}X_1)^H, \dots, (R_{g*}X_{k+1})^H) \\
&= \mathrm{d}\alpha_{p\cdot g}(R_{g*}(X_1^H), \dots, R_{g*}(X_{k+1}^H)) \\
&= (R_g^*\mathrm{d}\alpha)_p(X_1^H, \dots, X_{k+1}^H) \\
&= \mathrm{d}\big(R_g^*\alpha\big)_p(X_1^H, \dots, X_{k+1}^H) \\
&= \mathrm{Ad}_{g^{-1}}\mathrm{d}\alpha_p(X_1^H, \dots, X_{k+1}^H) \\
&= \mathrm{Ad}_{g_{-1}}\mathrm{d}^\omega\alpha_p(X_1, \dots, X_{k+1}).
\end{aligned}$$

He have then shown:

Lemma 2.8 (Exterior covariant derivative preserves basicness). *If $\alpha \in \Omega_{bas}(P, \mathfrak{g})$, then $\mathrm{d}^\omega\alpha \in \Omega^{k+1}_{bas}(P, \mathfrak{g})$.*

This result suggests that d^ω is particularly well-behaved on basic forms.

Proposition 2.9 (Expression for exterior covariant derivative on basic forms). *Let $\alpha \in \Omega^k_{bas}(P, \mathfrak{g})$ be a basic form. Then*

$$\mathrm{d}^\omega\alpha = \mathrm{d}\alpha + [\omega, \alpha].$$

Proof. — Let's consider the right-hand side. Let $X_0, \dots, X_k$ be vectors on T_pP. If all of them are horizontal, then the term $[\omega, \alpha]$ vanishes on them because, by definition, ω vanishes on horizontal vectors, and we end up with the definition of the exterior covariant derivative. Recalling the coordinate-free expression for the exterior differential

$$\begin{aligned}
\mathrm{d}\alpha(X_0, \dots, X_k) = &\sum_{j=0}^{k}(-1)^j X_j(\alpha(X_0, \dots, \hat{X}_j, \dots, X_k)) \\
&+ \sum_{i<j}(-1)^{i+j}\alpha([X_i, X_j], X_0, \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_k),
\end{aligned}$$

we see that the whole thing vanishes whenever there is more than 1 vertical vector, since we will always end up evaluating α in one of them. Similarly, we can see that in the evaluation of the bracket (following lemma 2.2),

$$[\omega, \alpha](X_0, \dots, X_k) = \frac{1}{k!}\sum_{\sigma\in\mathfrak{S}_{k+1}} \mathrm{sgn}(\sigma)[\omega(X_{\sigma(0)}), \alpha(X_{\sigma(1)}, \dots, X_{\sigma(k)})],$$

if there is more than one vertical vector, we will always evaluate α in one of them, so everything vanishes. Then, since $\mathrm{d}^\omega(\alpha)$ is horizontal, we trivially obtain the result.

The only non-trivial case is the one where we evaluate in exactly one vertical vector. Without loss of generality, suppose X_0 is vertical and $X_1, \dots, X_k$ are horizontal. We still have that

$$\mathrm{d}^\omega \alpha(X_0, \dots, X_k) = 0,$$

so we need to show that

$$\mathrm{d}\alpha(X_0, \dots, X_k) = -[\omega, \alpha](X_0, \dots, X_k).$$

On the right-hand side, we see that the evaluation of $[\omega, \alpha]$ reduces to the sum of the permutations where we evaluate ω on the vertical vector X_0, that is,

$$\begin{aligned}
[\omega, \alpha](X_0, \dots, X_k) &= \frac{1}{k!} \sum_{\substack{\sigma \in \mathfrak{S}_{k+1} \\ \sigma(0)=0}} \mathrm{sgn}(\sigma)[\omega(X_{\sigma(0)}), \alpha(X_{\sigma(1)}, \dots, X_{\sigma(k)})] \\
&= \frac{1}{k!} \sum_{\sigma' \in \mathfrak{S}_k} \mathrm{sgn}(\sigma')[\omega(X_0), \alpha(X_{\sigma'(1)}, \dots, X_{\sigma'(k)})] \\
&= \frac{1}{k!} \sum_{\sigma' \in \mathfrak{S}_k} \mathrm{sgn}(\sigma')^2[\omega(X_0), \alpha(X_1, \dots, X_k)] \\
&= [\omega(X_0), \alpha(X_1, \dots, X_k)].
\end{aligned}$$

Here we used the fact that a permutation that fixes 0 can be written as $\sigma(0) = 0; \sigma(i) = \sigma'(i)$ with $\sigma' \in \mathfrak{S}_k$, and these satisfy $\mathrm{sgn}(\sigma') = \mathrm{sgn}(\sigma)$. We have also used the fact that α is antisymmetric.

Now we want to evaluate $\mathrm{d}\alpha$, and for such we will use the long coordinate-free expression of the exterior derivative. First, letting $\xi = \omega_p(X_0) \in \mathfrak{g}$, we can extend X_0 to a vertical vector field (which we denote with the same symbol), as $X_0(p) = a_p(\xi)$; i.e. to the fundamental vector field associated to ξ. Second, we can also extend the vectors $X_1, \dots, X_k$ to horizontal vector fields that are furthermore *G-invariant*. To do so, we extend $\mathrm{d}_p\pi(X_j) \in T_{\pi(p)}M$ to a vector field on M, and consider its horizontal lift (see section 1.4), which we denote with the same symbol X_j. With this construction, since horizontal lifts are G-invariant and G-invariant fields commute with fundamental vector fields (lemma 1.12), we have that

$$\alpha([X_i, X_j], X_0, \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_k) = 0.$$

This follows for $i = 0$, since we evaluate on the bracket of a fundamental vector field and a G-invariant field, which is vanishing. When $i > 0$, we are evaluating α directly on a vertical field, so everything vanishes as well. Then we need only consider

$$\mathrm{d}\alpha(X_0, \dots, X_k) = \sum_{j=0}^{k} (-1)^j X_j(\alpha(X_0, \dots, \hat{X}_j, \dots, X_k)) = X_0(\alpha(X_1, \dots, X_k)).$$

The only term in the sum that does not immediately vanish is the one where we don't evaluate α on X_0. Now we evaluate at a point p. An integral curve of X_0 at p is $t \mapsto p \cdot \exp(t\xi)$, and we write $g_t = \exp(t\xi)$, so

$$\begin{aligned}
\mathrm{d}\alpha_p(X_0, \dots, X_k) &= X_0(p)(\alpha(X_1, \dots, X_k)) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \alpha_{p \cdot g_t}(X_1(p \cdot g_t), \dots, X_k(p \cdot g_t)) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \alpha_{p \cdot g_t}(\mathrm{d}_p R_{g_t}(X_1(p)), \dots, \mathrm{d}_p R_{g_t}(X_k(p))) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (R_{g_t}^* \alpha)_p(X_1(p), \dots, X_k(p)) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{Ad}_{g_t^{-1}} \alpha_p(X_1(p), \dots, X_k(p)) \\
&= \mathrm{ad}(-\xi)(\alpha_p(X_1(p), \dots, X_k(p))) \\
&= -[\xi, \alpha_p(X_1(p), \dots, X_k(p))] \\
&= -[\omega(X_0), \alpha_p(X_1(p), \dots, X_k(p))].
\end{aligned}$$

■

A corollary of this expression is that d^ω is not nilpotent. This means that we cannot (immediately) construct a cohomology theory based on basic forms and the exterior covariant derivative!

Corollary 2.10 (Exterior covariant derivative is not nilpotent). *Let $\varphi \in \Omega^0_{bas}(P, \mathfrak{g})$. Then*

$$(\mathrm{d}^\omega \circ \mathrm{d}^\omega)\varphi = [\Omega, \varphi].$$

Proof. — We have

$$\begin{aligned}
\mathrm{d}^\omega(\mathrm{d}^\omega\varphi) &= \mathrm{d}(\mathrm{d}^\omega\varphi) + [\omega, \mathrm{d}^\omega\varphi] \\
&= \mathrm{d}(\mathrm{d}\varphi + [\omega,\varphi]) + [\omega, \mathrm{d}\varphi] + [\omega,[\omega,\varphi]] \\
&= \mathrm{d}[\omega,\varphi] + [\omega,\mathrm{d}\varphi] + [\omega[\omega,\varphi]] \\
&= [\mathrm{d}\omega\,,\varphi] - [\omega,\mathrm{d}\varphi] + [\omega,\mathrm{d}\varphi] + [\omega[\omega,\varphi]] \\
&= [\mathrm{d}\omega\,,\varphi] + [\omega,[\omega,\varphi]].
\end{aligned}$$

Here we used the fact that for $\alpha \in \Omega^k(P,\mathfrak{g})$ and $\beta \in \Omega^l(P,\mathfrak{g})$:

$$\mathrm{d}[\alpha,\beta] = [\mathrm{d}\alpha\,,\beta] + (-1)^k[\alpha,\mathrm{d}\beta].$$

This can be readily checked from the definition, and it follows since the bracket is defined in terms of the wedge product.
Now let's evaluate at two vectors $u, v \in TP$:

$$\begin{aligned}
[\omega,[\omega,\varphi]](u,v) &= [\omega(u),[\omega,\varphi](v)] - [\omega(v),[\omega,\varphi](u)] \\
&= [\omega(u),[\omega(v),\varphi]] - [\omega(v),[\omega(u),\varphi]] \\
&= -[\omega(u),[\varphi,\omega(v)]] - [\omega(v),[\omega(u),\varphi]] \\
&= [\varphi,[\omega(v),\omega(u)]] \\
&= [[\omega(u),\omega(v)],\varphi] \\
&= \left[\frac{1}{2}[\omega,\omega],\varphi\right](u,v).
\end{aligned}$$

Therefore, we obtain

$$\mathrm{d}^\omega(\mathrm{d}^\omega\varphi) = [\mathrm{d}\omega,\varphi] + \frac{1}{2}[[\omega,\omega],\varphi] = [\Omega,\varphi]. \qquad \blacksquare$$

3 The relation with connections on vector bundles

3.1 From vector bundles to principal bundles

Let's go back to known waters. Let $\pi_E : E \to M$ be a vector bundle of rank k over M. Recall that a **connection** ∇ on E is (at least in one of its several flavors) a bilinear map

$$\nabla : \mathfrak{X}(M) \times \Gamma(E) \to \Gamma(E),$$

where we denote $\nabla(X)(s) = \nabla_X(s)$, such that for all $X \in \mathfrak{X}(M)$, $s \in \Gamma(E)$ and $f \in C^\infty(M)$:

1. $\nabla_{fX}s = f\nabla_X s$, and
2. $\nabla_X(fs) = f\nabla_X s + \mathcal{L}_X(f)s$ (Leibniz rule).

At this point, we know that we have a special principal $\mathrm{GL}(k,\mathbb{R})$-bundle that is directly related to E, namely the frame bundle $\mathrm{Fr}(E)$. Is there any relation between the connection ∇ and possible connections on $\mathrm{Fr}(E)$? Can we find a connection 1-form $\omega_\nabla \in \Omega^1(\mathrm{Fr}(E),\mathfrak{gl}(k,\mathbb{R}))$ that is induced by ∇?
Indeed, we can. First, we can rethink this map by fixing $s \in \Gamma(E)$. With s held fixed, we can then write

$$\begin{aligned}
\nabla s : \mathfrak{X}(M) &\to \Gamma(E) \\
X &\mapsto \nabla_X s.
\end{aligned}$$

By property (1) above, the map ∇s is $C^\infty(M)$-linear, and so we can interpret it as an E-valued 1-form on M:

$$\nabla s \in \Omega^1(M,E).$$

If $f \in C^\infty(M)$ is a function, then from the Leibniz rule we obtain that for all $X \in \mathfrak{X}(M)$,

$$\nabla(fs)(X) = \nabla_X(fs) = (\mathcal{L}_X f)s + f\nabla_X s = \mathrm{d}f\,(X)s + f\nabla s(X),$$

so we may write

$$\nabla(fs) = \mathrm{d}f \otimes s + f\nabla s$$

Now let U be a trivializing open set of the bundle, and let $\{e_1, \dots e_k\}$ be a frame on $E_U := \pi^{-1}(U)$. Of course, each element e_j is a section of E, so we can consider $\nabla e_j \in \Omega^1(U, E_U)$ (why E_U and not just E?). In particular, we can write ∇e_j as

$$\nabla e_j = \sum_i \Gamma^i_j e_i,$$

where each $\Gamma^i_j \in \Omega^1(U)$ is a 1-form. We can collect all the Γ^i_j in a $\mathfrak{gl}(k, \mathbb{R})$-valued form, whose entries are called the **connection coefficients** (or in some cases, the Christoffel symbols)

$$\Gamma = \begin{pmatrix} \Gamma^1_1 & \dots & \Gamma^1_k \\ \vdots & \ddots & \vdots \\ \Gamma^k_1 & \dots & \Gamma^k_k \end{pmatrix} \in \Omega^1(U, \mathfrak{gl}(k, \mathbb{R})).$$

What do we have at this point? For each frame $\{e_1, \dots, e_k\}$ of E, which is defined locally on $U \subseteq M$, we have a $\mathfrak{gl}(k, \mathbb{R})$-valued 1-form Γ. This smells quite a lot like what we're looking for! If we can show that the connection coeffients transform nicely with respect to change of frames, we can invoke the physicists-did-nothing-wrong proposition (proposition 1.10) and construct a connection on $\mathrm{Fr}(E)$.

So let $e'_1, \dots, e'_k$ be another frame, defined on an open $U' \subseteq M$. On $U \cap U'$, each element e'_j can be expressed in terms of the first frame. For each $x \in U \cap U'$ there is a matrix $A(x) \in \mathrm{GL}(k, \mathbb{R})$ such that

$$e'_j(x) = \sum_i A(x)^i{}_j e_i(x),$$

or rather, we have a $\mathrm{GL}(k, \mathbb{R})$-valued function A on $U \cap U'$, which is precisely the transition function of the trivialization of $\mathrm{Fr}(E)$. Now, when we evaluate the connection on e'_j, we get

$$\begin{aligned}
\nabla e'_j &= \sum_i \nabla(A^i{}_j e_i) \\
&= \sum_i \left(\mathrm{d}A^i{}_j \otimes e_i + A^i{}_j \nabla e_i\right) \\
&= \sum_i \left(\mathrm{d}A^i{}_j \otimes e_i + A^i{}_j \sum_r \Gamma^r_i e_r\right) \\
&= \sum_i \left(\mathrm{d}A^i{}_j + \sum_r A^r{}_j \Gamma^i_r\right) \otimes e_i.
\end{aligned}$$

On the other hand,

$$\nabla e'_j = \sum_r \Gamma'^r_j e'_r = \sum_{i,r} \Gamma'^r_j A^i{}_r e_i.$$

Comparing with the previous result, we obtain

$$\sum_r \Gamma'^r_j A^i{}_r = \mathrm{d}A^i{}_j + \sum_r A^r{}_j \Gamma^i_r.$$

Noting that the upper index is the column index, we see that the previous equation is for the components of the matrix equation

$$A\Gamma' = \mathrm{d}A + \Gamma A,$$

that is

$$\Gamma' = A^{-1}\Gamma A + A^{-1}\,\mathrm{d}A\,.$$

Indeed, we can now invoke proposition 1.10 and claim:

Theorem 3.1 (Connection induced by connection on vector bundle). *Let ∇ be a connection on a vector bundle $E \to M$ of rank k. Then there is a unique connection 1-form ω_∇ on the frame bundle $\mathrm{Fr}(E)$ such that, given a local frame $e : U \to \mathrm{Fr}(E)$, the local gauge potential is given by the connection coefficients:*

$$e^*\omega_\nabla = \Gamma.$$

There's also a direct way to construct ω_∇ given a connection ∇, that does not require using the physicists-did-nothing-wrong proposition. It can be found in (Bär, 2011, example 2.3.3) and (Crainic, 2015, section 2.3.5).

3.2 From principal bundles to vector bundles

The converse can be done with a little bit more generality. let $G \hookrightarrow P \xrightarrow{\pi} M$ be a principal G-bundle, V a vector space and $\rho : G \to \mathrm{GL}(V)$ a representation. We can construct the associated bundle $E = E(P, V, \rho) = P \times_\rho V$, defined as the quotient of $P \times V$ under the action

$$(p, v) \cdot g = (p \cdot g, \rho(g^{-1})v).$$

We will denote $\rho(g)v$ simply as $g \cdot v$ whenever there is no chance for confusion[4], and the elements of E in terms of representatives, e.g. $[p, v]$. The associated bundle is a vector bundle with fiber V, so we now can ask ourselves if, given a connection ω on P, there is an induced connection ∇^ω on E.

As above, for any connection ∇ on E, given a section $s \in \Gamma(E)$, we have an E-valued 1-form $\nabla s \in \Omega^1(M, E)$, so we can think of a connection as a map $\nabla : \Gamma(E) \to \Omega^1(M, E)$. Now we dig up some an important isomorphism, which is that E-valued forms on M correspond to basic V-valued forms on P, i.e. we have isomorphisms

$$h : \Omega^k_{\text{bas}}(P, V) \xrightarrow{\sim} \Omega^k(M, E).$$

In this case, we say that a form $\alpha \in \Omega^k(P, V)$ is basic if

1. α is horizontal, i.e. $\iota_X \alpha = 0$ for any vertical vector $X \in TP$; and
2. α is pseudotensorial of type ρ, that is,
$$R_g^* \alpha = \rho(g^{-1}) \circ \alpha.$$

Noting that a section of E is just an E-valued 0-form, we see that the problem is reduced to finding a suggestively-named map

$$\mathrm{d}^\omega : \Omega^0_{\text{bas}}(P, V) \to \Omega^1_{\text{bas}}(P, V),$$

that is *nicely* related to ω and that satisfies the Leibniz rule when we go back to M. Once we have such a map, we can define ∇^ω on E such that the following diagram commutes:

$$\begin{tikzcd}
\Omega^0_{\text{bas}}(P, V) \arrow[r, "\mathrm{d}^\omega"] \arrow[d, "h"] & \Omega^1_{\text{bas}}(P, V) \arrow[d, "h"] \\
\Gamma(E) \arrow[r, "\nabla^\omega"] & \Omega^1(M, E)
\end{tikzcd}$$

But wait a minute... for the case where $\rho = \mathrm{Ad}$ and $V = \mathfrak{g}$, we already have a such a map, namely the exterior covariant derivative d^ω, which acts on basic forms according to proposition 2.9 as

$$\mathrm{d}^\omega \alpha = \mathrm{d}\alpha + [\omega, \alpha].$$

And now we use the ancient art of reverse-engineering. If α is a 0-form, we can rewrite $[\omega, \alpha]$ in terms of the adjoint representation, precisely as $[\omega, \alpha] = \mathrm{ad}(\omega)(\alpha)$, where $\mathrm{ad} = \mathrm{d}_e \mathrm{Ad}$, so that

$$\mathrm{d}^\omega \alpha = \mathrm{d}\alpha + (\mathrm{d}_e \mathrm{Ad} \circ \omega)(\alpha).$$

This suggests that for a general vector space V and representation $\rho : G \to \mathrm{GL}(V)$, we define

$$\mathrm{d}^\omega \alpha := \mathrm{d}\alpha + (\mathrm{d}_e \rho \circ \omega)(\alpha),$$

on all basic 0-forms. Explicitly, for $p \in P$ and $X \in T_p P$, it is defined as

$$\mathrm{d}^\omega \alpha_p(X) = \mathrm{d}_p \alpha(X) + (\mathrm{d}_e \rho)(\omega_p(X))(\alpha(p)).$$

What we now need to show is that the map

$$\nabla^\omega := h \circ \mathrm{d}^\omega \circ h^{-1} : \Gamma(E) \to \Omega^1(M, E),$$

satisfies the Leibniz rule,

$$\nabla^\omega (fs) = \mathrm{d}f \otimes s + f \nabla^\omega s.$$

for all $f \in C^\infty(M)$ and $s \in \Gamma(E)$.

[4] which would be about... never

To prove this, we need to get our hands dirty, and explicitly write the isomorphism. If $\alpha \in \Omega^k_{\mathrm{bas}}(P, V)$, then $h(\alpha) \in \Omega^k(M, E)$ is given on $X_1, \ldots, X_k \in T_x M$ as

$$h(\alpha)_x(X_1, \ldots, X_k) = [p, \alpha_p(\tilde{X}_1, \ldots, \tilde{X}_k)],$$

where $\pi(p) = x$ and $\mathrm{d}_p \tilde{X}_j = X_j$. For the inverse, we first recall that each $p \in P$ determines an isomorphism $i_p : V \to E_p$ as $i_p(v) = [p, v]$. Then if $\phi \in \Omega^k(M, E)$, we have that for $Y_1 \ldots, Y_k \in T_p P$, $h^{-1}(\phi) \in \Omega^k_{\mathrm{bas}}(P, V)$

$$h^{-1}(\phi)_p(Y_1, \ldots, Y_k) = i_p^{-1}(\phi_{\pi(p)}(\pi_* Y_1, \ldots, \pi_* Y_k)).$$

It is a standard straightforward exercise to show that *everything works.*

Now let $f \in C^\infty(M)$ and $s \in \Gamma(E)$. The basic 0-form induced on P by fs is

$$h^{-1}(fs)(p) = i_p^{-1}((f(\pi(p))s(\pi(p))) = (f \circ \pi)(p) i_p^{-1}(s(\pi(p))),$$

and so $h^{-1}(fs) = (f \circ \pi) h^{-1}(s)$. Write $\tilde{f} = f \circ \pi$, and $\tilde{s} = h^{-1}(s)$. Then $\tilde{f}$ is a G-invariant real-valued function and $\tilde{s}$ is a basic V-valued function. Now we apply d^ω:

$$\mathrm{d}^\omega(\tilde{f}\tilde{s}) = \mathrm{d}(\tilde{f}\tilde{s}) + (\mathrm{d}_e\rho \circ \omega)(\tilde{f}\tilde{s}) = \mathrm{d}\tilde{f}\,\tilde{s} + \tilde{f}\,\mathrm{d}\tilde{s} + \tilde{f}(\mathrm{d}_e\rho \circ \omega)(\tilde{s}) = \mathrm{d}\tilde{f}\,\tilde{s} + \tilde{f}\,\mathrm{d}^\omega \tilde{s}\,.$$

Here we have that $\tilde{f}$ comes out of the differential of the representation, because once evaluated at $\omega_p(X)$ for some $p \in P$, $X \in T_p P$, $(\mathrm{d}_e\rho)(\omega_p(X))$ is *linear*. Now we apply h, evaluate at a point $x \in M$ and a vector $X \in T_x M$:

$$\begin{aligned} \nabla^\omega(fs)_x(X) &= h(\mathrm{d}^\omega\big(h^{-1}(fs)\big))_x(X) \\ &= h(\mathrm{d}\tilde{f}\,\tilde{s} + \tilde{f}\,\mathrm{d}^\omega \tilde{s})_x(X) \\ &= [p, \mathrm{d}_p \tilde{f}(\tilde{X})\tilde{s}(p) + \tilde{f}(p)\,\mathrm{d}^\omega \tilde{s}_p\,(\tilde{X})]. \end{aligned}$$

Now we recall that $\tilde{f} = f \circ \pi$, so $\tilde{f}(p) = f(x)$ and

$$\mathrm{d}_p \tilde{f}(\tilde{X}) = \mathrm{d}_x f \,\mathrm{d}_p \pi(\tilde{X}) = \mathrm{d}_x f(X).$$

Therefore

$$\begin{aligned} \nabla^\omega(fs)_x(X) &= [p, \mathrm{d}_x f(X)\tilde{s}(p)] + [p, f(x)\,\mathrm{d}^\omega \tilde{s}_p\,(\tilde{X})] \\ &= \mathrm{d}_x f(X)[p, \tilde{s}(p)] + f(x)[p, \mathrm{d}^\omega \tilde{s}_p\,(\tilde{X})] \\ &= (\mathrm{d}f \otimes s + f \nabla^\omega s)_p(X). \end{aligned}$$

Then ∇^ω is, indeed, a connection on E.

References

Bär, C. (2011). *Lecture notes on Gauge Theories*.
Crainic, M. (2015). *Lecture notes for the Mastermath course Differential Geometry 2015/2016*.
Kobayashi, S. and K. Nomizu (1996). *Foundations of Differential Geometry*. Vol. 1. Wiley Classics Library. Wiley.
Lee, J. M. (2012). *Introduction to Smooth Manifolds*. Second edition. Springer.
Morita, S. (2001). *Geometry of Differential Forms*. American Mathematical Society.
Naber, G. L. (2010). *Topology, Geometry and Gauge Fields: Foundations*. Second edition. Springer.
Nicolaescu, L. I. (2018). *Lectures on the Geometry of Manifolds*.
Quintero Vélez, A. (2018). *Notas de Fundamentos Matemáticos de las Teorías de Campos Gauge*.

Last edited: April 9, 2019