Connections on principal bundles

Santiago Quintero de los Ríos*

December 16, 2020

Contents

1 Connections on principal bundles **1**
1.1 Connections as horizontal distributions . . . 1
1.2 Connections as 1-forms . . . 3
1.3 Local expressions, or, why physicists did nothing wrong . . . 5
1.4 Horizontal lifts, parallel transport and holonomy . . . 8

2 Curvature **11**
2.1 The curvature 2-form and structure equation . . . 11
2.2 Local expressions (Curvature Edition) . . . 14
2.3 The exterior covariant derivative . . . 15

3 The relation with connections on vector bundles **17**
3.1 From vector bundles to principal bundles . . . 17
3.2 Interlude: Associated bundles . . . 19
3.3 From principal bundles to vector bundles . . . 22
3.4 In physics language . . . 24

Notation

These notes compile some general facts about connections on principal bundles, and their relation to connections on vector bundles.

A few notes on notation: We are working in the context of a **principal G-bundle** P over a manifold M. This we denote as $G \hookrightarrow P \xrightarrow{\pi} M$; where $\pi : P \to M$ is the projection map. The right action of G on P is denoted as $\sigma : P \times G \to P$. For any $g \in G$, we denote right multiplication by g as $\sigma_g : P \to P$; and for every $p \in P$, we denote the orbit map as $\sigma_p : G \to P$.

Given a smooth function $f : M \to N$ between manifolds, we denote the tangent map at some $x \in M$ as $T_x f : T_x M \to T_{f(x)} N$. This is to explicitly show the functoriality of T_x.

1 Connections on principal bundles

1.1 Connections as horizontal distributions

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

$$T_p \pi(v) = 0.$$

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 $T_p \pi$, so we have a short exact sequence

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

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

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

$$T_pP \cong V_pP \oplus T_{\pi(p)}M.$$

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

$$T_pP = V_pP \oplus H_p.$$

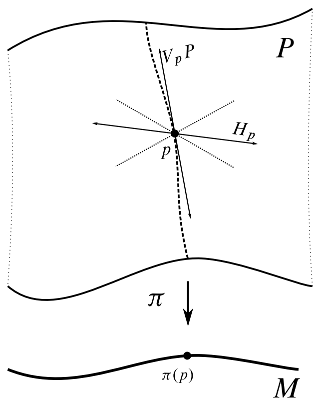

Figure 1: A choice of a horizontal space H_p at T_pP. There are many such choices (in dotted lines).

Once we have chosen a single horizontal subspace $H_p \subset T_pP$ 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 $\sigma_g(p) = p \cdot g$, is a fiber-preserving diffeomorphism, and thus $T_p\sigma_g$ is an isomorphism of tangent spaces that preserves the vertical subspace. This suggests that $T_p\sigma_g(H_p)$ is a horizontal subspace at $p \cdot g$. Indeed, noting that

$$T_{p\cdot g}\pi \circ T_p\sigma_g = T_p(\pi \circ \sigma_g) = T_p\pi(v),$$

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

$$u = T_p\sigma_g(T_{p\cdot g}\sigma_{g^{-1}}(u)) = T_p\sigma_g(\tilde{u}),$$

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

$$V_{p\cdot g}P = T_p\sigma_g(V_pP).$$

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

$$T_{p\cdot g}P = T_p\sigma_g(T_pP) = T_p\sigma_g(V_pP) \oplus T_p\sigma_g(H_p) = V_{p\cdot g}P \oplus T_p\sigma_g(H_p),$$

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 $g \in G$, $T_p\sigma_g(H_p)$ 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** *or* **Ehresmann 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 $p \in P$,*

$$T_p\sigma_g(H_p) = H_{p\cdot g}.$$

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:

$$a_p(X) = T_e\sigma_p(X).$$

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 $a_p = T_e\sigma_p$ 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 $g \in G$,

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

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 $\mathrm{Ad}_g = T_e\,\mathrm{Conj}_g$. Then we have

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

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 $X \in \mathfrak{g}$:

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

[1] https://xkcd.com/927/

and so we conclude that

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

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

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 $T_p \pi(v^V) = 0$, 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

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

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

$$T_p \sigma_g(H_p) = H_{p \cdot g}.$$

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*.

[2] We can handwave it away by saying that it follows from the smoothness of the distribution H.

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

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

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

$$\Theta_g = T_g L_{g^{-1}}.$$

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

$$a_g(X) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} g\exp(tX) = T_e L_g(X),$$

so that

$$\Theta_g(a_g(X)) = T_g(L_{g^{-1}})(T_e L_g(X)) = T_g(L_{g^{-1}} \circ L_g)(X) = X.$$

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

$$(\sigma_g^* \Theta)_h(X) = \Theta_{hg}(T_h \sigma_g(X)) = T_{hg} L_{g^{-1}h^{-1}} T_h \sigma_g(X) = T_h(L_{g^{-1}h^{-1}} \circ \sigma_g)(X).$$

But then, we see that

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

such that the differential at h is

$$T_h(L_{g^{-1}h^{-1}} \circ \sigma_g) = T_h(\mathrm{Conj}_{g^{-1}} \circ L_{h^{-1}}) = T_e \,\mathrm{Conj}_{g^{-1}}\, T_h L_{h^{-1}} = \mathrm{Ad}_{g^{-1}} \circ \Theta_h,$$

and so, indeed

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

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 $H_{(x,g)} := T_x M \oplus 0 \subset T_{(x,g)}(M \times G)$.

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 figure 2.

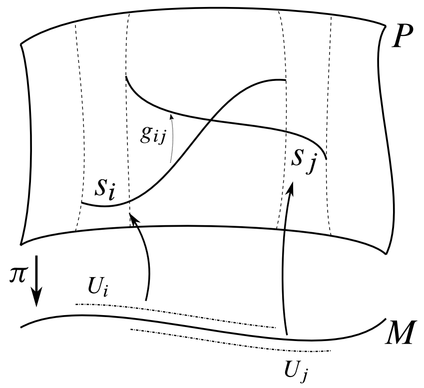

Figure 2: The transition functions g_{ij} relate the sections s_i, s_j induced by the trivializations.

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

$$\mathcal{A}_i := s_i^* \omega$$

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

Proposition 1.8 (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 $\mathcal{A}_i = s_i^*\omega$ be the local gauge potentials. Then for all $x \in U_{ij} = U_i \cap U_j$,

$$(\mathcal{A}_j)_x = \mathrm{Ad}_{g_{ij}(x)^{-1}} \circ (\mathcal{A}_i)_x + (g_{ij}^* \Theta)_x, \tag{1}$$

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

We write this compactly as

$$\mathcal{A}_j = \mathrm{Ad}_{g_{ij}^{-1}} \mathcal{A}_i + g_{ij}^* \Theta.$$

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

$$(\mathcal{A}_j)_x = (s_j^* \omega)_x = \omega_{s_j(x)} \circ T_x s_j,$$

so we need to find the expression for $T_x s_j$, 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

$$T_x s_j = T_x(\sigma \circ (s_i, g_{ij})) = T_{(s_j(x), g_{ij}(x))}\sigma \circ T_x(s_j, g_{ij}) = T_{(s_j(x), g_{ij}(x))}\sigma \circ (T_x s_j, T_x g_{ij}).$$

Now we need to find the expression for $T_{(p,g)}\sigma$. 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

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

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

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

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

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

We put these two together, and obtain

$$T_{(p,g)}\sigma(u,\xi) = T_p\sigma_g(u) + a_{p\cdot g}(\Theta_g(\xi)).$$

Substituting in T_xs_j, and evaluating at some $v \in T_xU_{ij}$,

$$\begin{aligned}
T_xs_j(v) = T_{(s_i(x),g_{ij}(x))}\sigma(T_xs_j(v), T_xg_{ij}(v)) &= T_{s_i(x)}\sigma_{g_{ij}(x)}(T_xs_i(v)) + a_{s_i(x)\cdot g_{ij}(x)}(\Theta_{g_{ij}(x)}(T_xg_{ij}(v))). \\
&= T_{s_i(x)}\sigma_{g_{ij}(x)}(T_xs_i(v)) + a_{s_j(x)}((g_{ij}^*\Theta)_x(v)).
\end{aligned}$$

Now we evaluate $\omega_{s_j(x)}$ on $T_xs_j(v)$. 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

$$\begin{aligned}
\omega_{s_j}(T_xs_j(u)) &= \omega_{s_j}(T_{s_i}\sigma_{g_{ij}}(T_xs_i(u))) \\
&= \omega_{s_ig_{ij}}(T_{s_i}\sigma_{g_{ij}}(T_xs_i(u))) \\
&= (\sigma_{g_{ij}}^*\omega)_{s_i}(T_xs_i(u)) \\
&= \mathrm{Ad}_{g_{ij}^{-1}}(\omega_{s_i}(T_xs_i(u))) \\
&= \mathrm{Ad}_{g_{ij}^{-1}}((s_i^*\omega)_x(u)) \\
&= (\mathrm{Ad}_{g_{ij}(x)^{-1}} \circ (\mathcal{A}_i)_x)(u).
\end{aligned}$$

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

In the previous proof we calculated the differential of the group action $\sigma : P \times G \to P$. We will use it a bit more so let's collect it in a lemma.

> **Lemma 1.9 (Differential of the group action).**
> *Let $P \to M$ be a principal G-bundle and denote by $\sigma : P \times G \to P$ the right action. Its differential is given by*
>
> $$T_{(p,g)}\sigma(u,\xi) = T_p\sigma_g(u) + a_{p\cdot g}(\Theta_g(\xi)),$$
>
> *for all $u \in T_pP$ and $\xi \in T_gG$. Here, $\sigma_g : P \to P$ is right multiplication by g, a_p is the infinitesimal action on p, and Θ is the Maurer-Cartan form.*

If the Lie group G is a *matrix* Lie group, then this result takes a particularly simple form. In a matrix Lie group, the adjoint representation is simply

$$\mathrm{Ad}_g(X) = gXg^{-1}.$$

The pullback of the Maurer-Cartan form also has a simple form. Let $X \in T_xM$ be a tangent vector with integral curve γ. Then if $g : U \subseteq M \to G$ is a smooth map,

$$\begin{aligned}(g^*\Theta)_x(X) &= \Theta_{g(x)}(T_x g(X)) \\ &= T_{g(x)} L_{g(x)^{-1}} T_x g(X) \\ &= T_x(L_{g(x)^{-1}} \circ g)(X) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g(x)^{-1} g(\gamma(t)) \\ &= g(x)^{-1} \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} g(\gamma(t)) \\ &= g(x)^{-1}(\mathrm{d}g)_x(X).\end{aligned}$$

Therefore, the gauge transformation of the gauge potential for a matrix Lie group is

$$\mathcal{A}_j = g_{ij}^{-1}\mathcal{A}_i g_{ij} + g_{ij}^{-1}\,\mathrm{d}g_{ij}\,.$$

This proposition, in physics, is often called *gauge transformation* of a potential. In physics we mostly work 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 equation (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 $\mathcal{A}_i \in \Omega^1(U_i, \mathfrak{g})$, such that for all $x \in U_i \cap U_j$,

$$(\mathcal{A}_j)_x = \mathrm{Ad}_{g_{ij}(x)^{-1}} \circ (\mathcal{A}_i)_x + g_{ij}^*\Theta_x.$$

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

$$s_i^*\omega = \mathcal{A}_i.$$

1.4 Horizontal lifts, parallel transport and holonomy

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 $T_p\pi(Y) = X$. 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 and $H \subset TP$ an Ehresmann connection on P. We define the **horizontal lift** *of X as the vector field $Y \in \mathfrak{X}(P)$, which satisfies $\pi_*Y = X$ and $Y_p \in H_p$ for all $p \in P$.*

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

$$\sigma_{g*}Y = Y.$$

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

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

Now suppose that we have a curve $\gamma : [0,1] \to M$. At each point over the curve, we have a vector $\dot\gamma(t) \in T_{\gamma(t)}M$, which we can lift we can lift to the fiber above $\gamma(t)$. So if we choose a starting point $p_0 \in \pi^{-1}(\gamma(0))$, we can find an integral curve along all these lifted vectors on the fibers over the curve γ. In the end we obtain a curve $\tilde\gamma : [0,1] \to P$ which satisfies $\pi \circ \tilde\gamma = \gamma$, $\tilde\gamma(0) = p_0$, and $\dot{\tilde\gamma}(t) \in H_{\tilde\gamma(t)}$ for all t. We call it a **horizontal lift** of γ, and it is unique:

Proposition 1.13 (Existence and uniqueness of horizontal lifts of curves). *Let ω be a connection on the G-bundle $P \to M$, and $\gamma : [0,1] \to M$ a piecewise smooth curve. Given a point $p \in \pi^{-1}(\gamma(0))$, there exists a* unique *curve $\tilde\gamma : [0,1] \to P$, called the* **horizontal lift** *of γ, satisfying*

1. *$\tilde\gamma$ is a* lift *of γ: $\pi \circ \tilde\gamma = \gamma$.*
2. *$\tilde\gamma$ is* horizontal*: $\dot{\tilde\gamma}(t) \in H_{\tilde\gamma(t)}$ for all $t \in [0,1]$.*
3. *$\tilde\gamma(0) = p$.*

Proof. — There's two ways to prove this: The first way is in the spirit of the discussion above. We have a vector field $\tilde X$ on the bundle $P|_{\gamma([0,1])}$,[3] where for $p \in \pi^{-1}(\gamma(t))$, $\tilde X_p$ is the horizontal lift of $\dot\gamma(t)$ to p. Then $\tilde\gamma$ is the integral curve of $\tilde X$ starting at the prescribed $p_0 \in \pi^{-1}(\gamma(0))$. Technically these integral curves only exist locally but since $[0,1]$ is compact we can glue a finite number of them together and be done.

The second way follows [Bär11, Lemma 2.6.1], where we look at a local problem in terms of sections and an ODE. We'll go through it because y'all know I love me some local descriptions of things.

Suppose that the image $\gamma([0,1])$ is contained in a single open set U that trivializes the bundle. In general the image is compact, so it will be contained in a union of finitely many of these. So there is an associated section $s : U \to P$. Any lift $\tilde\gamma$ will be of the form

$$\tilde\gamma(t) = s(\gamma(t)) \cdot g(t),$$

for some unique map $g[0,1] \to G$. The condition for $\tilde\gamma$ to be a *horizontal* lift is that $\dot{\tilde\gamma}(t) \in \ker \omega_{\tilde\gamma(t)}$ for all t.

First, we need to see what $\dot{\tilde\gamma}$ is. For simplicity, write $p(t) := s(\gamma(t))$.

$$\begin{aligned}
\dot{\tilde\gamma}(t) &= \frac{\mathrm{d}}{\mathrm{d}t} p(t)\cdot g(t) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\sigma(p(t), g(t)) \\
&= T_{(p(t),g(t))}\sigma\,(\dot p(t), \dot g(t)) \\
&= T_{p(t)}\sigma_{g(t)}(\dot p(t)) + a_{p(t)\cdot g(t)}(\Theta_{g(t)}(\dot g(t)).
\end{aligned}$$

Here we used the differential of σ from lemma 1.9. Now we apply $\omega_{\tilde g(t)}$. Recall that $\omega_p(a_p(\xi)) = \xi$ by definition, so

$$\begin{aligned}
\omega_{\tilde\gamma(t)}(\dot{\tilde\gamma}(t)) &= \omega_{p(t)\cdot g(t)}\left(T_{p(t)}\sigma_{g(t)}(\dot p(t)) + a_{p(t)\cdot g(t)}(\Theta_{g(t)}(\dot g(t))\right) \\
&= \omega_{p(t)\cdot g(t)}\left(T_{p(t)}\sigma_{g(t)}(\dot p(t))\right) + \Theta_{g(t)}(\dot g(t)) \\
&= (\sigma^*_{g(t)}\omega)_{p(t)}(\dot p(t)) + \Theta_{g(t)}(\dot g(t)) \\
&= \mathrm{Ad}_{g(t)^{-1}}\,\omega_{p(t)}(\dot p(t)) + \Theta_{g(t)}(\dot g(t)).
\end{aligned}$$

[3] We could talk of the pullback bundle $\gamma^*P \to [0,1]$ to be even *more* technical.

Now recall that Ad_g is the derivative at e of the conjugation map $\mathrm{Conj}_g(h) = ghg^{-1}$, which is precisely $L_g \circ R_{g^{-1}}$. Also recall that $\Theta_g = T_g L_{g^{-1}}$. These two things, put together give us

$$\begin{aligned}\omega_{\tilde{\gamma}(t)}(\dot{\tilde{\gamma}}(t)) &= T_e(L_{g(t)^{-1}} \circ R_{g(t)})\omega_{p(t)}(\dot{p}(t)) + T_{g(t)}L_{g(t)^{-1}}(\dot{g}(t)) \\ &= T_{g(t)}L_{g(t)^{-1}}\left(T_e R_{g(t)}\omega_{p(t)}(\dot{p}(t)) + \dot{g}(t)\right).\end{aligned}$$

Since $\tilde{\gamma}$ is *horizontal*, then this must be precisely zero. But since $T_{g(t)}L_{g(t)^{-1}}$ is an isomorphism (because left multiplication is a diffeomorphism), then the condition for $\tilde{\gamma}$ to be a horizontal lift is

$$T_e R_{g(t)}\omega_{p(t)}(\dot{p}(t)) + \dot{g}(t) = 0.$$

Finally, let's rewrite $p(t) = (s \circ \gamma)(t)$. Then $\dot{p}(t) = T_{\gamma(t)}s(\dot{\gamma}(t))$, and this becomes

$$T_e R_{g(t)}(s^*\omega)_{\gamma(t)}(\dot{\gamma}(t)) + \dot{g}(t) = 0.$$

This is a first order ordinary differential equation for $g(t)$, with a given initial condition, so it determines $g(t)$ uniquely. ■

Note that in terms of the local potential $\mathcal{A} = s^*\omega$, the local condition for $\tilde{\gamma(t)}$ to be a horizontal lift is

$$T_e R_{g(t)}\mathcal{A}_{\gamma(t)}(\dot{\gamma}(t)) + \dot{g}(t) = 0.$$

Even more, if G is a matrix Lie group, then right multiplication is a linear map so this becomes

$$\dot{g}(t) = -\mathcal{A}_{\gamma(t)}(\dot{\gamma}(t))g(t).$$

If the exponential map is surjective, then $g(t) = \exp(\xi(t))$, and this equation becomes

$$\dot{\xi}(t) = -\mathcal{A}_{\gamma(t)}(\dot{\gamma}(t)),$$

which has a solution

$$\xi(t) = \xi(0) - \int_0^t \mathcal{A}_{\gamma(\tau)}(\dot{\gamma}(\tau))\,\mathrm{d}\tau = \xi(0) - \int_\gamma \mathcal{A}.$$

Given a curve $\gamma : [0,1] \to M$, if we write $x_0 = \gamma(0)$ and $x_1 = \gamma(1)$, then we have a map $\mathrm{PT}_\gamma : \pi^{-1}(x_0) \to \pi^{-1}(x_1)$, called **parallel transport** along γ, where for $p \in \pi^{-1}(x_0)$, its image $\mathrm{PT}_\gamma(p)$ is the endpoint of the horizontal lift of γ with initial value p.

Note that if $\tilde{\gamma}$ is the horizontal lift starting at p, then for any $g \in G$, $\tilde{\gamma} \cdot g$ is a horizontal lift starting at $p \cdot g$. This tells us that

$$\mathrm{PT}_\gamma(p \cdot g) = \mathrm{PT}_\gamma(p) \cdot g.$$

In particular, since the action of G is transitive on the fibers of the bundle, then PT is necessarily bijective. Furthermore, if we choose a "reference point" $p \in \pi^{-1}(x_0)$, we have an isomorphism $\pi^{-1}(x_0) \cong G$ by $p \cdot g \mapsto g$ and similarly for $\pi^{-1}(x_1)$. Then under these isomorphisms, parallel transport $\mathrm{PT}_\gamma : \pi^{-1}(x_0) \to \pi^{-1}(x_1)$ is a "group isomorphism".

More specifically, suppose that γ is a loop based at x_0. Then parallel transport is an isomorphism $\mathrm{PT}_\gamma : \pi^{-1}(x_0) \to \pi^{-1}(x_0)$, and it determines a unique map $g_\gamma : \pi^{-1}(x_0) \to G$ which satisfies

$$\mathrm{PT}_\gamma(p) := p \cdot g_\gamma(p).$$

On one hand, we have for any $h \in G$,

$$\mathrm{PT}_\gamma(p \cdot g) = (p \cdot h) \cdot g_\gamma(p \cdot h),$$

but on the other hand

$$\mathrm{PT}_\gamma(p \cdot g) = \mathrm{PT}_\gamma(p) \cdot h,$$

which implies that the map g_γ satisfies

$$g_\gamma(p \cdot h) = h^{-1}g(p)h.$$

This means that the loop γ determines a *conjugation class* of G, as

$$\gamma \mapsto g_\gamma(\pi^{-1}x_0) = \left\{h^{-1}g_\gamma(p)h \,:\, h \in G\right\}.$$

2 Curvature

2.1 The curvature 2-form and structure equation

Let $G \hookrightarrow P \xrightarrow{\pi} M$ 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 $\alpha \wedge \beta$, $[\alpha \wedge \beta]$, or $\alpha \wedge_{[,]} \beta$...).

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 $X_1, \ldots, X_{i+j}$:

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

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 $G \hookrightarrow P \xrightarrow{\pi} M$. The **curvature** of ω is a $\mathfrak{g}$-valued 2-form $\Omega \in \Omega^2(P, \mathfrak{g})$ defined as*

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

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

$$\Theta_g = T_g(L_{g^{-1}}).$$

The Maurer-Cartan form is also left-invariant,

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

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;

$$\xi_a(g) := T_e L_g(e_a),$$

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

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

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

Proposition 2.5 (Curvature is basic).
Let ω be a connection on $G \hookrightarrow P \xrightarrow{\pi} M$ and Ω its curvature. Then $\Omega \in \Omega^2_{bas}(P, \mathfrak{g})$, that is,

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

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

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

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

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

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

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 $\mathrm{Ad}_g = T_e C_g$ 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

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

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

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

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

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

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

$$T_{p\cdot g_t}\sigma_{g_t^{-1}}(\nu_{p\cdot g_t}) \in H_p \qquad \text{for all } t,$$

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 2.6.
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 Local expressions (Curvature Edition)

Again, let's see what the curvature looks like once we take trivializations. Let $\{U_i\}_{i\in I}$ be a trivializing cover with local gauges $s_i : U_i \to P$, and with gauge transformations $g_{ij} : U_i \cap U_j \to G$, satisfying $s_j = s_i \cdot g_{ij}$. Let $\omega \in \Omega^1(P, \mathfrak{g})$ be a connection on P with curvature $\Omega \in \Omega^2(P, \mathfrak{g})$. For each local gauge, define the **gauge field strengths** $\mathcal{F}_i \in \Omega^2(U_i, \mathfrak{g})$ as

$$\mathcal{F}_i := s_i^*\Omega.$$

From the Cartan structure equation (equation (2)), we immediately obtain

$$\mathcal{F}_i = \mathrm{d}\mathcal{A}_i + \frac{1}{2}[\mathcal{A}_i, \mathcal{A}_i],$$

where $\mathcal{A}_i = s_i^*\omega$ are the local gauge potentials.

Again, how do these relate to one another? Let $X, Y \in T_xM$ be tangent vectors. Then, since the curvature Ω is a horizontal form,

$$\mathcal{F}_{j,x}(X, Y) = \Omega_{s_j(x)}(T_xs_j(X), T_xs_j(Y)) = \Omega_{s_j(x)}(T_xs_j(X)^H, T_xs_j(Y)^H).$$

In the proof of proposition 1.8, we showed that the differential T_xs_j is

$$T_xs_j(X) = T_{s_i(x)}\sigma_{g_{ij}(x)}(T_xs_i(X)) + a_{s_j(x)}((g_{ij}^*\Theta)_x(X)).$$

Note that the second term in this expression is a *vertical* vector, since it is in the image of the infinitesimal Lie group action. Therefore,

$$\begin{aligned}
\mathcal{F}_{j,x}(X, Y) &= \Omega_{s_j(x)}(T_{s_i(x)}\sigma_{g_{ij}(x)}(T_xs_i(X))^H, T_{s_i(x)}\sigma_{g_{ij}(x)}(T_xs_i(Y))^H) \\
&= \Omega_{s_i(x)g_{ij}(x)}(T_{s_i(x)}\sigma_{g_{ij}(x)}(T_xs_i(X)), T_{s_i(x)}\sigma_{g_{ij}(x)}(T_xs_i(Y))) \\
&= (\sigma_{g_{ij}(x)}^*\Omega)_{s_i(x)}(T_xs_i(X), T_xs_i(Y)) \\
&= \mathrm{Ad}_{g_{ij}(x)^{-1}}(\Omega_{s_i(x)}(T_xs_i(X), T_xs_i(Y))) \\
&= \mathrm{Ad}_{g_{ij}(x)^{-1}}((s_i^*\Omega)_x(X, Y)) \\
&= \mathrm{Ad}_{g_{ij}(x)^{-1}}(\mathcal{F}_{i,x}(X, Y)).
\end{aligned}$$

We have proved the following:

Proposition 2.7 (Transformation of local field strenghts).
Let ω be a connection on $G \hookrightarrow P \xrightarrow{\pi} M$ with curvature Ω, 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 $\mathcal{F}_i = s_i^*\Omega$ be the local gauge field strenghts. Then for all $x \in U_{ij} = U_i \cap U_j$,

$$\mathcal{F}_{j,x} = \mathrm{Ad}_{g_{ij}(x)^{-1}} \circ \mathcal{F}_{i,x}. \tag{3}$$

We write this compactly as

$$\mathcal{F}_j = \mathrm{Ad}_{g_{ij}^{-1}} \mathcal{F}_i.$$

2.3 The exterior covariant derivative

From corollary 2.6, 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 σ_g preserves horizontality and the pullback commutes with d, so in general pulling back by σ_g should behave reasonable well. Indeed, let $\alpha \in \Omega^k(P,\mathfrak{g})$ be pseudotensorial of type Ad. Then

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

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),

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

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,

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

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 $\operatorname{sgn}(\sigma')=\operatorname{sgn}(\sigma)$. 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 $T_p\pi(X_j)\in T_{\pi(p)}M$ 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

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

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

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

■

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

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

■

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 $GL(k,\mathbb{R})$-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

$$\nabla e_j = \sum_i \Gamma_j^i e_i,$$

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

$$\Gamma = \begin{pmatrix} \Gamma_1^1 & \dots & \Gamma_k^1 \\ \vdots & \ddots & \vdots \\ \Gamma_1^k & \dots & \Gamma_k^k \end{pmatrix} \in \Omega^1(U, \mathfrak{gl}(k,\mathbb{R})).$$

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 $A(x) \in GL(k,\mathbb{R})$ such that

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

or rather, we have a $GL(k,\mathbb{R})$-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

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

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 [Bär11, example 2.3.3] and [Cra15, section 2.3.5].

3.2 Interlude: Associated bundles

The converse of theorem 3.1 can be done with a little bit more generality, without any additional complications. We will see that given any connection $\omega \in \Omega^1(P, \mathfrak{g})$, we can find a connection on a wide array of vector bundles that are related to P.

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 define the **associated bundle** $P \times_\rho V$ 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, and the elements of E in terms of representatives, e.g. $[p, v]$. We have a projection map $\pi_E : E \to M$ a $\pi_E([p, v]) = \pi(p)$. This map is well-defined since $\pi(p \cdot g) = \pi(p)$ for all $p \in P$ and $g \in G$.

Of course, we have just given the definition as a set. We should check that $P \times_\rho V$ is 1. a manifold 2. a vector bundle. Usually we would skip this part but actually the construction of the charts and trivialization on E will give us a better understanding of it, and will tell us how it looks locally.

First, let's look at trivializations. Once we have the trivializations, we can construct a coordinate atlas adapted from the atlas of M, as with all fiber bundles.

For each $p \in P$, we define a map $i_p : V \to E$ as

$$i_p(v) = [p, v].$$

This map is a bijection from V to the fiber $E_{\pi(p)}$ above $\pi(p)$. Clearly, by construction i_p is surjective. And to see injectivity, suppose that $i_p(v) = i_p(v')$. Then $[p, v] = [p, v']$, so there exists a $g \in G$ such that $(p, v) = (p \cdot g, g^{-1} \cdot v')$. But the action of G is free on P, so necessarily $g = e$, and so $v = v'$. This allows us to endow $E_{\pi(p)}$ with a vector space structure such that i_p is a linear isomorphism; i.e. as $[p, v] + \alpha[p, v'] = [p, v + \alpha v']$ with the *same* p. Of course, this should be the same regardless of the choice of p (in the fiber of π, that is). Indeed, if $p' \in P$ is such that $\pi(p') = \pi(p)$, then there is a $g \in G$ such that $p' = p \cdot g$.

Therefore, we have that

$$i_{p'}(v) = [p', v] = [p \cdot g, v] = [p, g \cdot v] = i_p(g \cdot v) = (i_p \circ \rho(g))(v).$$

Then $i_{p'}$ and i_p are related by an automorphism of V, so the induced vector space structure on $E_{\pi(p)}$ is the same. This last result will be useful again later, so let's put it as a lemma.

Lemma 3.2.
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. Write* $E = P \times_\rho V$ *for the bundle associated to* P *via* ρ*. For* $p \in P$*, define* $i_p : V \to E_{\pi(p)}$ *as* $i_p(v) = [p, v]$*. Then* i_p *is a bijection, and for all* $g \in G$*,*

$$i_{p \cdot g} = i_p \circ \rho(g).$$

Let $\{U_i\}_{i \in I}$ be a cover of M that trivializes P. For each U_i we have a canonical section (or local gauge) $s_i : U_i \to P_{U_i}$ (see section 2.2). With this canonical section we can construct a trivialization $\Psi_i : U_i \times V \to E_{U_i}$ as

$$\Psi_i(x, v) = i_{s_i(x)}(v) = [s_i(x), v].$$

It can be shown [see e.g. Nab10, pp. 381] that when we endow E with the quotient topology, E is Hausdoff and each map Ψ_i is a homeomorphism. This is a straightforward (albeit a bit tedious) check.

How do the transition functions look? Consider two trivializing open sets U_i, U_j with their canonical sections s_i, s_j, and let $U_{ij} = U_i \cap U_j$. We have that there is a transition function $g_{ij} : U_{ij} \to G$ such that for all $x \in U_{ij}$,

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

Then, for $(x, v) \in U_{ij} \times V$, we have

$$\Psi_j(x, v) = [s_j(x), v] = [s_i(x) \cdot g_{ij}(x), v] = [s_i(x), g_{ij}(x) \cdot v] = \Psi_i(x, g_{ij}(v)).$$

This implies that

$$(\Psi_i^{-1} \circ \Psi_j)(x, v) = (x, g_{ij}(x) \cdot v) = (x, \rho(g_{ij}(x))(v)). \tag{4}$$

Then the transition functions are of the form $\rho(g_{ij}) \in \mathrm{GL}(V)$, with g_{ij} the gauge transitions of the principal bundle. This tells us that there is a unique smooth structure on E such that the Ψ_i are diffeomorphisms, and such that $\pi_E : E \to M$ is a smooth surjection. Thus, E is a vector bundle over M with typical fiber V. Let's put it as a proposition.

Proposition 3.3 (Associated bundle is a smooth vector bundle).
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. Write* $E = P \times_\rho V$ *for the bundle associated to* P *via* ρ*. Then* E *is a smooth vector bundle over* M *with typical fiber* V*. Furthermore, given a cover* $\{U_i\}_{i \in I}$ *that trivializes* P*, with canonical sections* $s_i : U_i \to P$ *and transition functions* $g_{ij} : U_i \cap U_j \to G$*, the maps* $\Psi_i : U_i \times V \to E$ *given as* $\Psi_i(x, v) = [s_i(x), v]$ *are trivializations of* E*, with transition functions* $\rho(g_{ij}) : U_i \cap U_j \to \mathrm{GL}(V)$*.*

In physics, we usually keep to local gauges. In a local gauge (U_i, s_i) , the bundle $P \times_\rho V$ is trivial and "looks like" $U_i \times V$. Equation (4) says that under a change of gauge $s_i \mapsto s_j$, an element $v \in V$ *transforms as* $v \to \rho(g_{ij})v$.

Example 3.4 (Frame bundles).
Let $E \xrightarrow{\pi_E} M$ be a real vector bundle of rank k over a smooth manifold M. We have a principal $\mathrm{GL}(k, \mathbb{R})$-bundle over M, which is the frame bundle $\mathrm{Fr}(E)$, and the identity representation of $\mathrm{GL}(k, \mathbb{R})$ on $\mathbb{R}^k$, $\mathrm{id} : \mathrm{GL}(k, \mathbb{R}) \to \mathrm{GL}(\mathbb{R}^k)$. It is a straightforward check to see that

$$E \cong \mathrm{Fr}(E) \times_{\mathrm{id}} \mathbb{R}^k.$$

Example 3.5 (Adjoint bundle).
Let $G \hookrightarrow P \xrightarrow{\pi} M$ be a principal G bundle. We have a natural representation $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$. The vector bundle associated to P via Ad is called the **adjoint bundle** of P, and is denoted $\mathrm{Ad}(P) := P \times_{\mathrm{Ad}} \mathfrak{g}$.

In particular, if $E \xrightarrow{\pi_E} M$ is a vector bundle of rank k, and $\mathrm{Fr}(E)$ is its frame bundle, then we have that

$$\mathrm{Ad}(\mathrm{Fr}(E)) \cong \mathrm{End}(E).$$

How do sections of E look like locally? If we fix a cover $\{U_i\}_{i\in I}$ of M that trivializes P, with canonical sections $s_i : U_i \to P$, we have that this cover also trivializes E. Suppose that $\psi : M \to E$ is a section of E. When restricted to U_i, we have that ψ looks locally like

$$(\Psi_i^{-1} \circ \psi)(x) = (x, \psi_i(x)),$$

for some smooth $\psi_i : U_i \to V$. In fact, given a local gauge s_i, there is a *bijection* between smooth maps $\psi_i : U_i \to V$ and local sections $\psi : U_i \to E$. On the overlaps $U_i \cap U_j$, the same argument of before shows that the "trivializations" of the sections transform as

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

Now we see that there is a deeper relation between sections of E (in fact, of E-valued forms) and V-valued forms on P. Again, let $\psi : M \to E$ be a section. In a local trivialization U_i with canonical section $s_i : U_i \to P$, we can write any element $p \in P_{U_i} = \pi^{-1}(U_i)$ uniquely as

$$p = s_i(x) \cdot g_i$$

where $x = \pi(p)$. By the discussion above, there is a function $\psi_i : U_i \to V$ such that for all $x \in U_i$, ψ looks like

$$\psi(x) = [s_i(x), \psi_i(x)].$$

So in general, we can change the representative of $\psi(x)$ to be of the form $[p, v]$ for any p in the fiber above x;

$$\psi(x) = [s_i(x), \psi_i(x)] = [s_i(x) \cdot g_i, g_i^{-1} \cdot \psi_i(x)].$$

We can thus define a function $\tilde{\psi}_i : P_{U_i} \to V$ as

$$\tilde{\psi}_i(s_i(x) \cdot g_i) := g_i^{-1} \cdot \psi_i(x).$$

It is a straightforward check to see that this function is well-defined on P_{U_i}; and in fact that the collection of $\{\tilde{\psi}_i\}_{i\in I}$ glues together to a map $\tilde{\psi} : P \to V$ which satisfies that for all $p \in P$ and $g \in G$,

$$\tilde{\psi}(p \cdot g) = g^{-1} \cdot \tilde{\psi}(p).$$

We say that $\tilde{\psi}$ is G-**equivariant** or **pseudotensorial of type** ρ.

This is a general fact: k-forms on M which are valued in E correspond to a certain kind to k-forms on P which are valued in P.

Definition 3.6 (Tensorial form).
Let $G \hookrightarrow P \xrightarrow{\pi} M$ be a principal G-bundle. We say that a form $\alpha \in \Omega^k(P, V)$ is **tensorial** *of type ρ or* **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,*

$$\sigma_g^* \alpha = \rho(g^{-1}) \circ \alpha.$$

We denote the space of tensorial k-forms of type ρ by $\Omega_\rho^k(P, V)$.

Theorem 3.7 (Lowering of tensorial forms).
Let $G \hookrightarrow P \xrightarrow{\pi} M$ be a principal bundle, V a vector space, $\rho : G \to \mathrm{GL}(V)$ a representation of G on V and $E = P \times_\rho V$ the associated vector bundle. Then there is a linear isomorphism $h : \Omega_\rho^k(P, V) \to \Omega^k(M, E)$.

Proof (Sketch). — Define $h : \Omega_\rho^k(P, V) \to \Omega^k(M, E)$ as follows: given $\overline{\phi} \in \Omega_\rho^k(P, V)$, define

$$h(\overline{\phi})_x(V_1, \dots, V_k) := \left[p, \overline{\phi}_p(\overline{V}_1, \dots, \overline{V}_k)\right],$$

where $p \in \pi^{-1}(x)$, $V_1, \dots, V_k \in T_xM$, and the $\overline{V}_i$ are lifts of the V_i; that is, $T_p\pi(\overline{V}_i) = V_i$ for $i = 1, 2, \dots, k$.

It is straightforward, but a bit long, to check that h is well-defined.

The inverse of h can be given explicitly: given $\psi \in \Omega^k(M, E)$, we define $h^{-1}\phi \in \Omega^k_\rho(P, V)$ as

$$(h^{-1}\phi)_p(\overline{V}_1, \dots, \overline{V}_k) := i_p^{-1}(\pi^*\phi)_x(\overline{V}_1, \dots, \overline{V}_k).$$

Again, it is a straightforward check to see that these maps are well-defined, linear, and inverses of one another. These maps are natural in the sense that they are the obvious choice given the data that we have. ∎

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.

3.3 From principal bundles to vector bundles

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 $\nabla : \Gamma(E) \to \Omega^1(M, E)$. Noting that a section of E is just an E-valued 0-form, and using the isomorphism h of theorem 3.7,we see that the problem is reduced to finding a suggestively-named map

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

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:

$$\begin{array}{ccc} \Omega^0_\rho(P, V) & \xrightarrow{\mathrm{d}^\omega} & \Omega^1_\rho(P, V) \\ \downarrow h & & \downarrow h \\ \Gamma(E) & \xrightarrow{\nabla^\omega} & \Omega^1(M, E) \end{array}$$

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 $\mathrm{ad} = T_e \mathrm{Ad}$, so that

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

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

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

on all basic 0-forms. The derivative $T_e\rho : \mathfrak{g} \to \mathrm{End}(V)$ is called the **infinitesimal action** of $\mathfrak{g}$ on V, induced by the action ρ.

Explicitly, for $p \in P$ and $X \in T_pP$, it is defined as

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

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 $s \in \Gamma(E)$.

To prove this, we need to get our hands dirty. 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:

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

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

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

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

$$T_p\tilde{f}(\tilde{X}) = T_xf\,T_p\pi(\tilde{X}) = T_xf(X).$$

Therefore

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

Then ∇^ω is, indeed, a connection on E. We have then proved

Theorem 3.8 (Connection induced by a connection on principal bundle).
Let $G \hookrightarrow P \xrightarrow{\pi} M$ be a principal G-bundle, V a vector space, $\rho : G \to \mathrm{GL}(V)$ a representation and $E = P \times_\rho V$ the associated bundle. Given a connection $\omega \in \Gamma^1(P, \mathfrak{g})$, there is an induced connection $\nabla^\omega : \Gamma(E) \to \Omega^1(M, E)$ such that the following diagram commutes:

$$\begin{CD} \Omega^0_\rho(P,V) @>{\mathrm{d}^\omega}>> \Omega^1_\rho(P,V) \\ @V{h}VV @VV{h}V \\ \Gamma(E) @>{\nabla^\omega}>> \Omega^1(M,E) \end{CD}\;.$$

Another way to prove this theorem is to go local, and define the connection in terms of the connection coefficients. This suffices to uniquely define a connection on a vector bundle [see Nic18, Proposition 3.3.5], if the coefficients transform well enough.

So let's go local, and try to see what the beast of $\nabla^\omega = h \circ \mathrm{d}^\omega \circ h^{-1}$ is. As always, let $\{U_i\}_{i\in I}$ be a cover of M that trivializes P, with canonical sections $s_i : U_i \to P_{U_i}$, and gauge transitions $g_{ij} : U_i \cap U_j \to G$. As we saw in section 3.2, this trivialization also induces a trivialization of the associated bundle $E = P \times_\rho V$.

We can go further and see that the trivialization of P also makes a further identification in the isomorphism $\Omega^k_\rho(P,V) \cong \Omega^k(M,E)$. Given a form $\alpha \in \Omega^k_\rho(P,V)$, we have that

$$h(\alpha)_x(V_1, \dots, V_k) = [p, \alpha_p(\overline{V}_1, \dots, \overline{V}_k)]$$

for $p \in \pi^{-1}(x)$ and $\overline{V}_1, \dots, \overline{V_k}$ lifts of the $V_1, \dots, V_k$. But we have a preferred choice of element in the fiber of x, namely $p = s_i(x)$. Similarly, we have a preferred lift of the V_j, namely as $\overline{V_j} = T_xs_i(V_j)$. We then have that

$$h(\alpha)_x(V_1, \dots, V_k) = [s_i(x), (s_i^*\alpha)_x(V_1, \dots, V_k)].$$

Thus, we have that the following diagram commutes:

$$\begin{array}{ccc} \Omega^k_\rho(P_{U_i}, V) & \xrightarrow{\;s_i^*\;} & \Omega^k(U_i, E) \\ & \searrow^{h} & \big\downarrow{\scriptstyle i_{s_i}} \\ & & \Omega^k(U_i, V) \end{array}\;.$$

This tells us that, given a choice of trivialization (gauge), V-valued tensorial forms on P and E-valued forms on M both reduce to V-valued forms on M.

In particular, given section $\psi : M \to E$, which looks locally on U_i as

$$\psi(x) = [s_i(x), \psi_i(x)]$$

the above diagram tells us that

$$\psi(x) = (h \circ h^{-1})(\psi(x))) = (i_{s_i(x)} \circ s_i^*)(h^{-1}(\psi(x))) = [s_i(x), s_i^*(h^{-1}\psi)(x)],$$

but since $i_{s_i(x)}$ is an isomorphism, then

$$\psi_i(x) = s_i^*(h^{-1}\psi)(x).$$

Now apply $\nabla^\omega \psi$. By the previous result,

$$\nabla^\omega \psi(x) = h(\mathrm{d}^\omega(h^{-1}\psi)(x)) = [s_i(x), s_i^*(\mathrm{d}^\omega h^{-1}\psi)(x)].$$

Therefore, it suffices to find $s_i^*(\mathrm{d}^\omega h^{-1}\psi)(x)$. For the first term, we have

$$s_i^*(\mathrm{d}h^{-1}\psi) = \mathrm{d}s_i^*(h^{-1}\psi) = \mathrm{d}\psi_i.$$

For the second term we need to be more careful. Let's evaluate at $x \in U_i$ and $V \in T_xM$:

$$s_i^*(T_e\rho \circ \omega(h^{-1}\psi))_x(V) = (T_e\rho(\omega_{s_i(x)}(s_{i*}V)))(h^{-1}\psi(s_i(x))) := (T_e\rho \circ \omega_i)_x(V)(\psi_i(x)).$$

Here we have denoted $\omega_i = s_i^*\omega$ (in consistence with the notation of proposition 1.8). Therefore, the local expression of the connection on E is (dropping the arguments)

$$\nabla^\omega \psi = [s_i, \mathrm{d}\psi_i + (T_e\rho \circ \omega_i)(\psi_i)].$$

3.4 In physics language

We can reduce the notation a bit more (and make it a bit more confusing), add coordinates, and obtain the equations of the "covariant derivative on matter fields" that is used by physicists. In physics, a matter field is a (local expression of a) section of the associated bundle E to a principal G-bundle. The group G is called the group of local invariance. A section $s : U \to P$ is a local gauge, and the local potentials of a connection are denoted by $\mathcal{A} := s^*\omega$.

The infinitesimal action is not denoted explicitly, so we only write $\xi \cdot v$ instead of $T_e\rho(\xi)(v)$, for $\xi \in \mathfrak{g}$ and $v \in V$. Thus, if we choose a basis $\{e_1, \dots, e_k\}$ of V, then we can write the infinitesimal action as a matrix product:

$$\mathcal{A} \cdot e_a = \mathcal{A}^b{}_a e_b,$$

with $\mathcal{A}^b{}_a \in \Omega^1(U)$ for all a, b. If we have a section $\Psi : M \to E$, then in the local gauge s it can be written as $\Psi = [s, \psi]$, with $\psi : U \to V$. This ψ is called a **matter field**, and in the basis of V, it becomes

$$\psi = \psi^a e_a.$$

If we assume that out trivializing cover is also a coordinate atlas, then on the chart U_i we also have coordinates x^μ. Therefore, we write

$$\mathcal{A}^b{}_a = \mathcal{A}^b{}_{a\mu}\, \mathrm{d}x^\mu.$$

Finally, when we apply the connection to Ψ

$$\nabla^\omega \Psi = [s, \mathcal{D}\psi],$$

where

$$\begin{aligned} \mathcal{D}\psi &= \mathrm{d}\psi + \mathcal{A} \cdot \psi \\ &= (\partial_\mu \psi^a + \mathcal{A}^a{}_{b\mu}\psi^b)\, \mathrm{d}x^\mu \otimes e_a \\ &:= \mathcal{D}_\mu \psi^a (\mathrm{d}x^\mu \otimes e_a). \end{aligned}$$

Here, the operator $\mathcal{D}_\mu$ is called the covariant derivative:

$$\mathcal{D}_\mu = \partial_\mu + \mathcal{A}_\mu.$$

References

[Bär11] Christian Bär. *Lecture notes on Gauge Theories*. 2011.

[Cra15] Marius Crainic. *Lecture notes for the Mastermath course Differential Geometry 2015/2016*. 2015.

[KN96] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of Differential Geometry*. Vol. 1. Wiley Classics Library. Wiley, 1996.

[Lee12] John M. Lee. *Introduction to Smooth Manifolds*. Second edition. Springer, 2012.

[Mor01] Shigeyuki Morita. *Geometry of Differential Forms*. American Mathematical Society, 2001.

[Nab10] Gregory L. Naber. *Topology, Geometry and Gauge Fields: Foundations*. Second edition. Springer, 2010.

[Nic18] Liviu I. Nicolaescu. *Lectures on the Geometry of Manifolds*. 2018.

[Qui18] Alexander Quintero Vélez. *Notas de Fundamentos Matemáticos de las Teorías de Campos Gauge*. 2018.