A connection $\omega$ on a principal bundle $P(M, G)$ separates $T_uP$ into $H_uP \oplus V_uP$. Accordingly, a vector $X \in T_uP$ is decomposed as $X = X^H + X^V$ where $X^H \in H_uP$ and $X^V \in V_uP$.

*Definition 10.4.* Let $\phi \in \Omega^r(P) \otimes V$ and $X_1, \ldots, X_{r+1} \in T_uP$. The **covariant derivative** of $\phi$ is defined by

$$\mathrm{D}\phi(X_1, \ldots, X_{r+1}) \equiv \mathrm{d}_P\, \phi(X_1^H, \ldots, X_{r+1}^H) \tag{10.28}$$

where $\mathrm{d}_P\, \phi \equiv \mathrm{d}_P\, \phi^\alpha \otimes e_\alpha$.

### 10.3.2 Curvature

*Definition 10.5.* The **curvature two-form** $\Omega$ is the covariant derivative of the connection one-form $\omega$,

$$\Omega \equiv \mathrm{D}\omega \in \Omega^2(P) \otimes \mathfrak{g}. \tag{10.29}$$

*Proposition 10.2.* The curvature two-form satisfies (cf (10.3b))

$$R_a^*\Omega = a^{-1}\Omega a \qquad a \in G. \tag{10.30}$$

*Proof.* We first note that $(R_{a*}X)^H = R_{a*}(X^H)$ ($R_{a*}$ preserves the horizontal subspaces) and $\mathrm{d}_P R_a^* = R_a^* \mathrm{d}_P$, see (5.75). By definition we find

$$\begin{aligned}
R_a^*\Omega(X, Y) &= \Omega(R_{a*}X, R_{a*}Y) = \mathrm{d}_P\omega\,((R_{a*}X)^H, (R_{a*}Y)^H) \\
&= \mathrm{d}_P\omega\,(R_{a*}X^H, R_{a*}Y^H) = R_a^*\, \mathrm{d}_P\omega\,(X^H, Y^H) \\
&= \mathrm{d}_P R_a^*\omega\,(X^H, Y^H) \\
&= \mathrm{d}_P\,(a^{-1}\omega a)(X^H, Y^H) = a^{-1}\mathrm{d}_P\omega\,(X^H, Y^H)a \\
&= a^{-1}\Omega(X, Y)a
\end{aligned}$$

where we noted that $a$ is a constant element and hence $\mathrm{d}_P a = 0$. □

Take a $\mathfrak{g}$-valued $p$-form $\zeta = \zeta^\alpha \otimes T_\alpha$ and a $\mathfrak{g}$-valued $q$-form $\eta = \eta^\alpha \otimes T_\alpha$ where $\zeta^\alpha \in \Omega^p(P)$, $\eta^\alpha \in \Omega^q(P)$, and $\{T_\alpha\}$ is a basis of $\mathfrak{g}$. Define the commutator of $\zeta$ and $\eta$ by

$$\begin{aligned}
[\zeta, \eta] &\equiv \zeta \wedge \eta - (-1)^{pq}\eta \wedge \zeta \\
&= T_\alpha T_\beta \zeta^\alpha \wedge \eta^\beta - (-1)^{pq} T_\beta T_\alpha \eta^\beta \wedge \zeta^\alpha \\
&= [T_\alpha, T_\beta] \otimes \zeta^\alpha \wedge \eta^\beta = f_{\alpha\beta}{}^\gamma T_\gamma \otimes \zeta^\alpha \wedge \eta^\beta.
\end{aligned} \tag{10.31}$$

If we put $\zeta = \eta$ in (10.31), when $p$ and $q$ are odd, we have

$$[\zeta, \zeta] = 2\zeta \wedge \zeta = f_{\alpha\beta}{}^\gamma T_\gamma \otimes \zeta^\alpha \wedge \zeta^\beta.$$

*Lemma 10.2.* Let $X \in H_uP$ and $Y \in V_uP$. Then $[X, Y] \in H_uP$.

*Proof*. Let $Y$ be a vector field generated by $g(t)$, then

$$\mathcal{L}_Y X = [Y, X] = \lim_{t\to 0} t^{-1}(R_{g(t)*}X - X).$$

Since a connection satisfies $R_{g*}H_uP = H_{ug}P$, the vector $R_{g(t)*}X$ is horizontal and so is $[Y, X]$. □

*Theorem 10.3.* Let $X, Y \in T_uP$. Then $\Omega$ and $\omega$ satisfy **Cartan's structure equation**

$$\Omega(X, Y) = \mathrm{d}_P\omega(X, Y) + [\omega(X), \omega(Y)] \qquad (10.32a)$$

which is also written as

$$\Omega = \mathrm{d}_P\omega + \omega \wedge \omega. \qquad (10.32b)$$

*Proof*. We consider the following three cases separately:

(i) Let $X, Y \in H_uP$. Then $\omega(X) = \omega(Y) = 0$ by definition. From definition 10.5, we have $\Omega(X, Y) = \mathrm{d}_P\omega(X^H, Y^H) = \mathrm{d}_P\omega(X, Y)$, since $X = X^H$ and $= Y^H$.

(ii) Let $X \in H_uP$ and $Y \in V_uP$. Since $Y^H = 0$, we have $\Omega(X, Y) = 0$. We also have $\omega(X) = 0$. Thus, we need to prove $\mathrm{d}_P\omega(X, Y) = 0$. From (5.70), we obtain

$$\mathrm{d}_P\omega(X, Y) = X\omega(Y) - Y\omega(X) - \omega([X, Y]) = X\omega(Y) - \omega([X, Y]).$$

Since $Y \in V_uP$, there is an element $V \in \mathfrak{g}$ such that $Y = V^{\#}$. Then $\omega(Y) = V$ is constant, hence $X\omega(Y) = X \cdot V = 0$. From lemma 10.2, we have $[X, Y] \in H_uP$ so that $\omega([X, Y]) = 0$ and we find $\mathrm{d}_P\omega(X, Y) = 0$.

(iii) For $X, Y \in V_uP$, we have $\Omega(X, Y) = 0$. We find that, in this case,

$$\mathrm{d}_P\omega(X, Y) = X\omega(Y) - Y\omega(X) - \omega([X, Y]) = -\omega([X, Y]).$$

We note that $X$ and $Y$ are closed under the Lie bracket, $[X, Y] \in V_uP$, see exercise 10.1(b). Then there exists $A \in \mathfrak{g}$ such that

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

where $A^{\#} = [X, Y]$. Let $B^{\#} = X$ and $C^{\#} = Y$. Then $[\omega(X), \omega(Y)] = [B, C] = A$ since $[B, C]^{\#} = [B^{\#}, C^{\#}]$. Thus, we have shown that

$$0 = \mathrm{d}_P\omega(X, Y) + \omega([X, Y]) = \mathrm{d}_P\omega(X, Y) + [\omega(X), \omega(Y)].$$

Since $\Omega$ is linear and skew symmetric, these three cases are sufficient to show that (10.32) is true for any vectors.

To derive (10.32b) from (10.32a), we note that

$$\begin{aligned}
[\omega, \omega](X, Y) &= [T_\alpha, T_\beta]\omega^\alpha \wedge \omega^\beta(X, Y) \\
&= [T_\alpha, T_\beta][\omega^\alpha(X)\omega^\beta(Y) - \omega^\beta(X)\omega^\alpha(Y)] \\
&= [\omega(X), \omega(Y)] - [\omega(Y), \omega(X)] = 2[\omega(X), \omega(Y)].
\end{aligned}$$

Hence, $\Omega(X, Y) = (\mathrm{d}_P\omega + \frac{1}{2}[\omega, \omega])(X, Y) = (\mathrm{d}_P\omega + \omega \wedge \omega)(X, Y)$. □

### 10.3.3 Geometrical meaning of the curvature and the Ambrose–Singer theorem

We have shown in chapter 7 that the Riemann curvature tensor expresses the non-commutativity of the parallel transport of vectors. There is a similar interpretation of curvature on principal bundles. We first show that $\Omega(X, Y)$ yields the vertical component of the Lie bracket $[X, Y]$ of horizontal vectors $X, Y \in H_u P$. It follows from $\omega(X) = \omega(Y) = 0$ that

$$\mathrm{d}_P\omega\,(X, Y) = X\omega(Y) - Y\omega(X) - \omega([X, Y]) = -\omega([X, Y]).$$

Since $X^H = X$, $Y^H = Y$, we have

$$\Omega(X, Y) = \mathrm{d}_P\omega\,(X, Y) = -\omega([X, Y]). \tag{10.33}$$

Let us consider a coordinate system $\{x^\mu\}$ on a chart $U$. Let $V = \partial/\partial x^1$ and $W = \partial/\partial x^2$. Take an infinitesimal parallelogram $\gamma$ whose corners are O $= \{0, 0, \ldots, 0\}$, P $= \{\varepsilon, 0, \ldots, 0\}$, Q $= \{\varepsilon, \delta, 0, \ldots, 0\}$ and R $= \{0, \delta, 0, \ldots, 0\}$. Consider the horizontal lift $\tilde{\gamma}$ of $\gamma$. Let $X, Y \in H_u P$ such that $\pi_* X = \varepsilon V$ and $\pi_* Y = \delta W$. Then

$$\pi_*([X, Y]^H) = \epsilon\delta[V, W] = \epsilon\delta\left[\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}\right] = 0 \tag{10.34}$$

that is $[X, Y]$ is *vertical*. This consideration shows that the horizontal lift $\tilde{\gamma}$ of a loop $\gamma$ fails to close. This failure is proportional to the vertical vector $[X, Y]$ connecting the initial point and the final point on the same fibre. The curvature measures this distance,

$$\Omega(X, Y) = -\omega([X, Y]) = A \tag{10.35}$$

where $A$ is an element of $\mathfrak{g}$ such that $[X, Y] = A^\#$.

Since the discrepancy between the initial and final points of the horizontal lift of a closed curve is simply the holonomy, we expect that the holonomy group is expressed in terms of the curvature.

*Theorem 10.4.* (**Ambrose–Singer theorem**) Let $P(M, G)$ be a $G$ bundle over a connected manifold $M$. The Lie algebra $\mathfrak{h}$ of the holonomy group $\Phi_{u_0}$ of a point $u_0 \in P$ agrees with the subalgebra of $\mathfrak{g}$ spanned by the elements of the form

$$\Omega_u(X, Y) \qquad X, Y \in H_u P \tag{10.36}$$

where $a \in P$ is a point on the same horizontal lift as $u_0$. [See Choquet-Bruhat *et al* (1982) for the proof.]

### 10.3.4 Local form of the curvature

The local form $\mathcal{F}$ of the curvature $\Omega$ is defined by

$$\mathcal{F} \equiv \sigma^*\Omega \tag{10.37}$$

where $\sigma$ is a local section defined on a chart $U$ of $M$ (cf $\mathcal{A} = \sigma^*\omega$). $\mathcal{F}$ is expressed in terms of the gauge potential $\mathcal{A}$ as

$$\mathcal{F} = \mathrm{d}\mathcal{A} + \mathcal{A} \wedge \mathcal{A} \tag{10.38a}$$

where d is the exterior derivative on $M$. The action of $\mathcal{F}$ on the vectors of $TM$ is given by

$$\mathcal{F}(X, Y) = \mathrm{d}\mathcal{A}\,(X, Y) + [\mathcal{A}(X), \mathcal{A}(Y)]. \tag{10.38b}$$

To prove (10.38a) we note that $\mathcal{A} = \sigma^*\omega$, $\sigma^*\,\mathrm{d}_P\omega = \mathrm{d}\sigma^*\omega$ and $\sigma^*(\zeta \wedge \eta) = \sigma^*\zeta \wedge \sigma^*\eta$. From Cartan's structure equation, we find

$$\mathcal{F} = \sigma^*(\mathrm{d}_P\omega + \omega \wedge \omega) = \mathrm{d}\sigma^*\omega + \sigma^*\omega \wedge \sigma^*\omega = \mathrm{d}\mathcal{A} + \mathcal{A} \wedge \mathcal{A}.$$

Next, we find the component expression of $\mathcal{F}$ on a chart $U$ whose coordinates are $x^\mu = \varphi(p)$. Let $\mathcal{A} = \mathcal{A}_\mu\,\mathrm{d}x^\mu$ be the gauge potential. If we write $\mathcal{F} = \frac{1}{2}\mathcal{F}_{\mu\nu}\,\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$, a direct computation yields

$$\mathcal{F}_{\mu\nu} = \partial_\mu\mathcal{A}_\nu - \partial_\nu\mathcal{A}_\mu + [\mathcal{A}_\mu, \mathcal{A}_\nu]. \tag{10.39}$$

$\mathcal{F}$ is also called the curvature two-form and is identified with the (**Yang–Mills**) **field strength**. To avoid confusion, we call $\Omega$ the curvature and $\mathcal{F}$ the (Yang–Mills) field strength. Since $\mathcal{A}_\mu$ and $\mathcal{F}_{\mu\nu}$ are $\mathfrak{g}$-valued functions, they can be expanded in terms of the basis $\{T_\alpha\}$ of $\mathfrak{g}$ as

$$\mathcal{A}_\mu = A_\mu{}^\alpha T_\alpha \qquad \mathcal{F}_{\mu\nu} = F_{\mu\nu}{}^\alpha T_\alpha. \tag{10.40}$$

The basis vectors satisfy the usual commutation relations $[T_\alpha, T_\beta] = f_{\alpha\beta}{}^\gamma T_\gamma$. We then obtain the well-known expression

$$F_{\mu\nu}{}^\alpha = \partial_\mu A_\nu{}^\alpha - \partial_\nu A_\mu{}^\alpha + f_{\beta\gamma}{}^\alpha A_\mu{}^\beta A_\nu{}^\gamma. \tag{10.41}$$

*Theorem 10.5.* Let $U_i$ and $U_j$ be overlapping charts of $M$ and let $\mathcal{F}_i$ and $\mathcal{F}_j$ be field strengths on the respective charts. On $U_i \cap U_j$, they satisfy the compatibility condition,

$$\mathcal{F}_j = \mathrm{Ad}_{t_{ij}^{-1}}\,\mathcal{F}_i = t_{ij}^{-1}\mathcal{F}_i t_{ij} \tag{10.42}$$

where $t_{ij}$ is the transition function on $U_i \cap U_j$.

*Proof.* Introduce the corresponding gauge potentials $\mathcal{A}_i$ and $\mathcal{A}_j$,

$$\mathcal{F}_i = \mathrm{d}\mathcal{A}_i + \mathcal{A}_i \wedge \mathcal{A}_i \qquad \mathcal{F}_j = \mathrm{d}\mathcal{A}_j + \mathcal{A}_j \wedge \mathcal{A}_j.$$

Substituting $\mathcal{A}_j = t_{ij}^{-1}\mathcal{A}_i t_{ij} + t_{ij}^{-1}\mathrm{d}t_{ij}$ into $\mathcal{F}_j$, we verify that

$$\begin{aligned}
\mathcal{F}_j &= \mathrm{d}\,(t_{ij}^{-1}\mathcal{A}_i t_{ij} + t_{ij}^{-1}\,\mathrm{d}t_{ij}) \\
&\quad + (t_{ij}^{-1}\mathcal{A}_i t_{ij} + t_{ij}^{-1}\,\mathrm{d}t_{ij}) \wedge (t_{ij}^{-1}\mathcal{A}_i t_{ij} + t_{ij}^{-1}\,\mathrm{d}t_{ij}) \\
&= [-t_{ij}^{-1}\,\mathrm{d}t_{ij} \wedge t_{ij}^{-1}\mathcal{A}_i t_{ij} + t_{ij}^{-1}\,\mathrm{d}\mathcal{A}_i t_{ij} \\
&\quad - t_{ij}^{-1}\mathcal{A}_i \wedge \mathrm{d}t_{ij} - t_{ij}^{-1}\,\mathrm{d}t_{ij}\, t_{ij}^{-1} \wedge \mathrm{d}t_{ij}] \\
&\quad + [t_{ij}^{-1}\mathcal{A}_i \wedge \mathcal{A}_i t_{ij} + t_{ij}^{-1}\mathcal{A}_i \wedge \mathrm{d}t_{ij} \\
&\quad + t_{ij}^{-1}\,\mathrm{d}t_{ij}\, t_{ij}^{-1} \wedge \mathcal{A}_i t_{ij} + t_{ij}^{-1}\,\mathrm{d}t_{ij} \wedge t_{ij}^{-1}\,\mathrm{d}t_{ij}] \\
&= t_{ij}^{-1}(\mathrm{d}\mathcal{A}_i + \mathcal{A}_i \wedge \mathcal{A}_i)t_{ij} = t_{ij}^{-1}\mathcal{F}_i t_{ij}
\end{aligned}$$

where use has been made of the identity $\mathrm{d}t^{-1} = -t^{-1}\,\mathrm{d}t\, t^{-1}$. □

*Exercise 10.7.* The gauge potential $\mathcal{A}$ is called a **pure gauge** if $\mathcal{A}$ is written locally as $\mathcal{A} = g^{-1}\,\mathrm{d}g$. Show that the field strength $\mathcal{F}$ vanishes for a pure gauge $\mathcal{A}$. [It can be shown that the converse is also true. If $\mathcal{F} = 0$ on a chart $U$, the gauge potential may be expressed *locally* as $\mathcal{A} = g^{-1}\,\mathrm{d}g$.]

### 10.3.5 The Bianchi identity

Since $\omega$ and $\Omega$ are $\mathfrak{g}$-valued, we expand them in terms of the basis $\{T_\alpha\}$ of $\mathfrak{g}$ as $\omega = \omega^\alpha T_\alpha$, $\Omega = \Omega^\alpha T_\alpha$. Then (10.32b) becomes

$$\Omega^\alpha = \mathrm{d}_P\omega^\alpha + f_{\beta\gamma}{}^\alpha \omega^\beta \wedge \omega^\gamma. \tag{10.43}$$

Exterior differentiation of (10.43) yields

$$\mathrm{d}_P\Omega^\alpha = f_{\beta\gamma}{}^\alpha\,\mathrm{d}_P\omega^\beta \wedge \omega^\gamma + f_{\beta\gamma}{}^\alpha \omega^\beta \wedge \mathrm{d}_P\omega^\gamma. \tag{10.44}$$

If we note that $\omega(X) = 0$ for a horizontal vector $X$, we find

$$\mathrm{D}\Omega(X, Y, Z) = \mathrm{d}_P\Omega\,(X^H, Y^H, Z^H) = 0$$

where $X, Y, Z \in T_u P$. Thus, we have proved the **Bianchi identity**

$$\mathrm{D}\Omega = 0. \tag{10.45}$$

Let us find the local form of the Bianchi identity. Operating with $\sigma^*$ on (10.44), we find that $\sigma^*\,\mathrm{d}_P\Omega = \mathrm{d}\cdot\sigma^*\Omega = \mathrm{d}\mathcal{F}$ for the LHS and

$$\begin{aligned}
\sigma^*(\mathrm{d}_P\omega \wedge \omega - \omega \wedge \mathrm{d}_P\omega) &= \mathrm{d}\sigma^*\omega \wedge \sigma^*\omega - \sigma^*\omega \wedge \mathrm{d}\sigma^*\omega \\
&= \mathrm{d}\mathcal{A} \wedge \mathcal{A} - \mathcal{A} \wedge \mathrm{d}\mathcal{A} = \mathcal{F} \wedge \mathcal{A} - \mathcal{A} \wedge \mathcal{F}
\end{aligned}$$

for the RHS. Thus, we have obtained that

$$\mathcal{D}\mathcal{F} = \mathrm{d}\mathcal{F} + \mathcal{A} \wedge \mathcal{F} - \mathcal{F} \wedge \mathcal{A} = \mathrm{d}\mathcal{F} + [\mathcal{A}, \mathcal{F}] = 0 \qquad (10.46)$$

where the action of $\mathcal{D}$ on a $\mathfrak{g}$-valued $p$-form $\eta$ on $M$ is defined by

$$\mathcal{D}\eta \equiv \mathrm{d}\eta + [\mathcal{A}, \eta]. \qquad (10.47)$$

Note that $\mathcal{D}\mathcal{F} = \mathrm{d}\mathcal{F}$ for $G = \mathrm{U}(1)$.

## 10.4 The covariant derivative on associated vector bundles

A connection one-form $\omega$ on a principal bundle $P(M, G)$ enables us to define the covariant derivative in associated bundles of $P$ in a natural way.

### 10.4.1 The covariant derivative on associated bundles

In physics, we often need to differentiate sections of a vector bundle which is associated with a certain principal bundle. For example, a charged scalar field in QED is regarded as a section of a complex line bundle associated with a U(1) bundle $P(M, \mathrm{U}(1))$. Differentiating sections covariantly is very important in constructing gauge-invariant actions.

Let $P(M, G)$ be a $G$ bundle with the projection $\pi_P$. Let us take a chart $U_i$ of $M$ and a section $\sigma_i$ over $U_i$. We take the canonical trivialization $\phi_i(p, e) = \sigma_i(p)$. Let $\tilde{\gamma}$ be a horizontal lift of a curve $\gamma : [0, 1] \to U_i$. We denote $\gamma(0) = p_0$ and $\tilde{\gamma}(0) = u_0$. Associated with $P$ is a vector bundle $E = P \times_\rho V$ with the projection $\pi_E$, see section 9.4. Let $X \in T_pM$ be a tangent vector to $\gamma(t)$ at $p_0$. Let $s \in \Gamma(M, E)$ be a section, or a vector field, on $M$. Write an element of $E$ as $[(u, v)] = \{(ug, \rho(g)^{-1}v | u \in P, v \in V, g \in G\}$. Taking a representative of the equivalence class amounts to fixing the gauge. We choose the following form,

$$s(p) = [(\sigma_i(p), \xi(p))] \qquad (10.48)$$

as a representative.

Now we define the parallel transport of a vector in $E$ along a curve $\gamma$ in $M$. Of course, a naive guess '$\xi$ is parallel transported if $\xi(\gamma(t))$ is constant along $\gamma(t)$' does not make sense since this statement depends on the choice of the section $\sigma_i(p)$. We define a vector to be parallel transported if it is constant with respect to a *horizontal lift* $\tilde{\gamma}$ of $\gamma$ in $P$. In other words, a section $s(\gamma(t)) = [(\tilde{\gamma}(t), \eta(\gamma(t)))]$ is parallel transported if $\eta$ is constant along $\gamma(t)$. This definition is intrinsic since if $\tilde{\gamma}'(t)$ is another horizontal lift of $\gamma$, then it can be written as $\tilde{\gamma}'(t) = \tilde{\gamma}(t)a$, $a \in G$ and we have (we omit $\rho$ to simplify the notation)

$$[(\tilde{\gamma}(t), \eta(t))] = [(\tilde{\gamma}'(t)a^{-1}, \eta(t))] = [(\tilde{\gamma}'(t), a^{-1}\eta(t))]$$

where $\eta(t)$ stands for $\eta(\gamma(t))$. Hence, if $\eta(t)$ is constant along $\gamma(t)$, so is its constant multiple $a^{-1}\eta(t)$.

Now the definition of covariant derivative is in order. Let $s(p)$ be a section of $E$. Along a curve $\gamma : [0, 1] \to M$ we have $s(t) = [(\tilde{\gamma}(t), \eta(t))]$, where $\tilde{\gamma}(t)$ is an arbitrary horizontal lift of $\gamma(t)$. The covariant derivative of $s(t)$ along $\gamma(t)$ at $p_0 = \gamma(0)$ is defined by

$$\nabla_X s \equiv \left[ \left( \tilde{\gamma}(0), \frac{\mathrm{d}}{\mathrm{d}t}\eta(\gamma(t))\bigg|_{t=0} \right) \right] \tag{10.49}$$

where $X$ is the tangent vector to $\gamma(t)$ at $p_0$. For the covariant derivative to be really intrinsic, it should not depend on the *extra* information, that is the special horizontal lift. Let $\tilde{\gamma}'(t) = \tilde{\gamma}(t)a$ $(a \in G)$ be another horizontal lift of $\gamma$. If $\tilde{\gamma}'(t)$ is chosen to be *the* horizontal lift, we have a representative $[(\tilde{\gamma}'(t), a^{-1}\eta(t))]$. The covariant derivative is now given by

$$\left[ \left( \tilde{\gamma}'(0), \frac{\mathrm{d}}{\mathrm{d}t}\{a^{-1}\eta(t)\}\bigg|_{t=0} \right) \right] = \left[ \left( \tilde{\gamma}'(0)a^{-1}, \frac{\mathrm{d}}{\mathrm{d}t}\eta(t)\bigg|_{t=0} \right) \right]$$

which agrees with (10.49). Hence, $\nabla_X s$ depends only on the tangent vector $X$ and the sections $s \in \Gamma(M, E)$ and not on the horizontal lift $\tilde{\gamma}(t)$. Our definition depends only on a curve $\gamma$ and a connection and not on local trivializations. The local form of the covariant derivative is useful in practical computations and will be given later.

So far we have defined the covariant derivative at a point $p_0 = \gamma(0)$. It is clear that if $X$ is a vector field, $\nabla_X$ maps a section $s$ to a new section $\nabla_X s$, hence $\nabla_X$ is regarded as a map $\Gamma(M, E) \to \Gamma(M, E)$. To be more precise, take $X \in \mathcal{X}(M)$ whose value at $p$ is $X_p \in T_pM$. There is a curve $\gamma(t)$ such that $\gamma(0) = p$ and its tangent at $p$ is $X_p$. Then any horizontal lift $\tilde{\gamma}(t)$ of $\gamma$ enables us to compute the covariant derivative $\nabla_X s|_p \equiv \nabla_{X_p} s$. We also define a map $\nabla : \Gamma(M, E) \to \Gamma(M, E) \otimes \Omega^1(M)$ by

$$\nabla s(X) \equiv \nabla_X s \qquad X \in \mathcal{X}(M) \qquad s \in \Gamma(M, E). \tag{10.50}$$

*Exercise 10.8.* Show that

$$\nabla_X(a_1 s_1 + a_2 s_2) = a_1 \nabla_X s_1 + a_2 \nabla_X s_2 \tag{10.51a}$$

$$\nabla(a_1 s_1 + a_2 s_2) = a_1 \nabla s_1 + a_2 \nabla s_2 \tag{10.51b}$$

$$\nabla_{(a_1 X_1 + a_2 X_2)} s = a_1 \nabla_{X_1} s + a_2 \nabla_{X_2} s \tag{10.51c}$$

$$\nabla_X(fs) = X[f]s + f\nabla_X s \tag{10.51d}$$

$$\nabla(fs) = (\mathrm{d}f)s + f\nabla s \tag{10.51e}$$

$$\nabla_{fX} s = f\nabla_X s \tag{10.51f}$$

where $a_i \in \mathbb{R}$, $s, s' \in \Gamma(M, E)$ and $f \in \mathcal{F}(M)$.

### 10.4.2 A local expression for the covariant derivative

In practical computations it is convenient to have a local coordinate representation of the covariant derivative. Let $P(M, G)$ be a $G$ bundle and $E = P \times_\rho G$ be an associate vector bundle. Take a local section $\sigma_i \in \Gamma(U_i, P)$ and employ the canonical trivialization $\sigma_i(p) = \phi_i(p, e)$. Let $\gamma : [0, 1] \to M$ be a curve in $U_i$ and $\tilde{\gamma}$ its horizontal lift, which is written as

$$\tilde{\gamma}(t) = \sigma_i(t) g_i(t) \tag{10.52}$$

where $g_i(t) \equiv g_i(\gamma(t)) \in G$. Take a section $e_\alpha(p) \equiv [(\sigma_i(p), e_\alpha{}^0)]$ of $E$, where $e_\alpha{}^0$ is the $\alpha$th basis vector of $V$; $(e_\alpha{}^0)^\beta = (\delta_\alpha)^\beta$. We have

$$e_\alpha(t) = [(\tilde{\gamma}(t) g_i(t)^{-1}, e_\alpha{}^0)] = [(\tilde{\gamma}(t), g_i(t)^{-1} e_\alpha{}^0)]. \tag{10.53}$$

Note that $g_i(t)^{-1}$ acts on $e_\alpha{}^0$ to compensate for the change of basis along $\gamma$. The covariant derivative of $e_\alpha$ is then given by

$$\begin{aligned}
\nabla_X e_\alpha &= \left[\left(\tilde{\gamma}(0), \frac{\mathrm{d}}{\mathrm{d}t}\{g_i(t)^{-1} e_\alpha{}^0\}\bigg|_{t=0}\right)\right] \\
&= \left[\left(\tilde{\gamma}(0), -g_i(t)^{-1}\left\{\frac{\mathrm{d}}{\mathrm{d}t} g_i(t)\right\} g_i(t)^{-1} e_\alpha{}^0\bigg|_{t=0}\right)\right] \\
&= [(\tilde{\gamma}(0) g_i(0)^{-1}, \mathcal{A}_i(X) e_\alpha{}^0)]
\end{aligned} \tag{10.54}$$

where (10.13b) has been used. From (10.54) we find the local expression,

$$\nabla_X e_\alpha = [(\sigma_i(0), \mathcal{A}_i(X) e_\alpha{}^0)]. \tag{10.55}$$

Let $\mathcal{A}_i = \mathcal{A}_{i\mu}\,\mathrm{d}x^\mu = \mathcal{A}_{i\mu}{}^\alpha{}_\beta\,\mathrm{d}x^\mu$ where $\mathcal{A}_{i\mu}{}^\alpha{}_\beta \equiv \mathcal{A}_{i\mu}{}^\gamma (T_\gamma)^\alpha{}_\beta$. The second entry of (10.55) is

$$\mathcal{A}_i(X) e_\alpha{}^0 = \frac{\mathrm{d}x^\mu}{\mathrm{d}t} e_\beta{}^0 \mathcal{A}_{i\mu}{}^\beta{}_\gamma \delta_\alpha{}^\gamma = \frac{\mathrm{d}x^\mu}{\mathrm{d}t} \mathcal{A}_{i\mu}{}^\beta{}_\alpha e_\beta{}^0.$$

Substituting this into (10.55), we finally have

$$\nabla_X e_\alpha = \left[\left(\sigma_i(0), \frac{\mathrm{d}x^\mu}{\mathrm{d}t} \mathcal{A}_{i\mu}{}^\beta{}_\alpha e_\beta{}^0\right)\right] = \frac{\mathrm{d}x^\mu}{\mathrm{d}t} \mathcal{A}_{i\mu}{}^\beta{}_\alpha e_\beta \tag{10.56a}$$

or

$$\nabla e_\alpha = \mathcal{A}_i{}^\beta{}_\alpha e_\beta. \tag{10.56b}$$

In particular, for a coordinate curve $x^\mu$, we have

$$\nabla_{\partial/\partial x^\mu} e_\alpha = \mathcal{A}_{i\mu}{}^\beta{}_\alpha e_\beta. \tag{10.57}$$

It is remarkable that a connection $\mathcal{A}$ on a principal bundle $P$ completely specifies the covariant derivative on an associated bundle $E$ (modulo representations).

*Exercise 10.9.* Let $s(p) = [(\sigma_i(p), \xi_i(p))] = \xi_i{}^\alpha(p)e_\alpha$ be a general section of $E$, where $\xi_i(p) = \xi_i{}^\alpha(p)e_\alpha{}^0$. Use the results of exercise 10.8 to verify that

$$\nabla_X s = \left[\left(\sigma_i(0), \frac{\mathrm{d}\xi_i}{\mathrm{d}t} + \mathcal{A}_i(X)\xi_i\bigg|_{t=0}\right)\right] = \frac{\mathrm{d}x^\mu}{\mathrm{d}t}\left\{\frac{\partial \xi_i{}^\alpha}{\partial x^\mu} + \mathcal{A}_{i\mu}{}^\alpha{}_\beta \xi_i{}^\beta\right\} e_\alpha. \tag{10.58}$$

By construction, the covariant derivative is independent of the local trivialization. This is also observed from the local form of $\nabla_X s$. Let $\sigma_i(p)$ and $\sigma_j(p)$ be local sections on overlapping charts $U_i$ and $U_j$. On $U_i \cap U_j$, we have $\sigma_j(p) = \sigma_i(p)t_{ij}(p)$. In the $i$-trivialization, the covariant derivative is

$$\begin{aligned}\nabla_X s &= \left[\left(\sigma_i(0), \frac{\mathrm{d}\xi_i}{\mathrm{d}t} + \mathcal{A}_i(X)\xi_i\bigg|_{t=0}\right)\right] \\ &= \left[\left(\sigma_j(0)\cdot t_{ij}^{-1}, \frac{\mathrm{d}}{\mathrm{d}t}(t_{ij}\xi_j) + \mathcal{A}_i(X)t_{ij}\xi_j\bigg|_{t=0}\right)\right] \\ &= \left[\left(\sigma_j(0), \frac{\mathrm{d}\xi_j}{\mathrm{d}t} + \mathcal{A}_j(X)\xi_j\bigg|_{t=0}\right)\right]\end{aligned} \tag{10.59}$$

where use has been made of the condition (10.9). The last line of (10.59) is $\nabla_X s$ expressed in the $j$-trivialization.

We have found that the covariant derivative defined by (10.49) is independent of the horizontal lift as well as the local section. The gauge potential $\mathcal{A}_i$ transforms under the change of local trivialization so that $\nabla_X s$ is a well-defined section of $E$. In this sense, $\nabla_X$ is the most natural derivative on an associated vector bundle, which is compatible with the connection on the principal bundle $P$.

*Example 10.4.* Let us recover the results obtained in section 7.2. Let $FM$ be a frame bundle over $M$ and let $TM$ be its associated bundle. We note $FM = P(M, \mathrm{GL}(m, \mathbb{R}))$ and $TM = FM \times_\rho \mathbb{R}^m$, where $m = \dim M$ and $\rho$ is the $m \times m$ matrix representation of $\mathrm{GL}(m, \mathbb{R})$. Elements of $\mathfrak{gl}(m, \mathbb{R})$ are $m \times m$ matrices. Let us rewrite the local connection form $\mathcal{A}_i$ as $\Gamma^\alpha{}_{\mu\beta}\,\mathrm{d}x^\mu$. We then find that

$$\nabla_{\partial/\partial x^\mu} e_\alpha = [(\sigma_i(0), \Gamma_\mu e_\alpha{}^0)] = \Gamma^\beta{}_{\mu\alpha} e_\beta \tag{10.60}$$

which should be compared with (7.14). For a general section (vector field), $s(p) = [(\sigma_i(p), X_i(p))] = X_i{}^\alpha(p)e_\alpha$, we find

$$\nabla_{\partial/\partial x^\mu} s = \left(\frac{\partial}{\partial x^\mu} X_i{}^\alpha + \Gamma^\alpha{}_{\mu\beta} X^\beta\right) e_\alpha \tag{10.61}$$

which reproduces the result of section 7.2. It is evident that the roles played by the indices $\alpha$, $\beta$ and $\mu$ in $\Gamma^\alpha{}_{\mu\beta}$ are very different in their characters; $\mu$ is the $\Omega^1(M)$ index while $\alpha$ and $\beta$ are the $\mathfrak{gl}(m, \mathbb{R})$ indices.

*Example 10.5.* Let us consider the U(1) gauge field coupled to a complex scalar field $\phi$. The relevant fibre bundles are the U(1) bundle $P(M, \mathrm{U}(1))$ and the associated bundle $E = P \times_\rho \mathbb{C}$ where $\rho$ is the natural identification of an element of U(1) with a complex number. The local expression for $\omega$ is $\mathcal{A}_i = \mathcal{A}_{i\mu}\, \mathrm{d}x^\mu$, where $\mathcal{A}_{i\mu} = \mathcal{A}_i(\partial/\partial x^\mu)$ is the vector potential of Maxwell's theory. Let $\gamma$ be a curve in $M$ with tangent vector $X$ at $\gamma(0)$. Take a local section $\sigma_i$ and express a horizontal lift $\tilde{\gamma}$ of $\gamma$ as $\tilde{\gamma}(t) = \sigma_i(t)\mathrm{e}^{\mathrm{i}\varphi(t)}$. If $1 \in \mathbb{C}$ is taken to be the basis vector, the basis section is

$$e = [(\sigma_i(p), 1)].$$

Let $\phi(p) = [(\sigma_i(p), \Phi(p))] = \Phi(p)e$ $(\Phi : M \to \mathbb{C})$ be a section of $E$, which is identified with a complex scalar field. With respect to $\tilde{\gamma}(t)$, the section is given by

$$\phi(t) = \Phi(t)[(\tilde{\gamma}(t), U(t)^{-1})] \tag{10.62}$$

where $U(t) = \mathrm{e}^{\mathrm{i}\varphi(t)}$. The covariant derivative of $\phi$ along $\gamma$ is

$$\begin{aligned}\nabla_X \phi &= \frac{\mathrm{d}\Phi}{\mathrm{d}t}[(\tilde{\gamma}(0), U(0)^{-1})] + \Phi(0)[(\tilde{\gamma}(0), U(0)^{-1}\mathcal{A}_i(X)\cdot 1)] \\ &= \left(\frac{\mathrm{d}\Phi}{\mathrm{d}t} + \mathcal{A}_{i\mu}\Phi\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\right)e = X^\mu\left(\frac{\partial\Phi}{\partial x^\mu} + \mathcal{A}_{i\mu}\Phi\right)e.\end{aligned} \tag{10.63}$$

*Example 10.6.* Let us consider the SU(2) Yang–Mills theory on $M$. The relevant bundles are the SU(2) bundle $P(M, \mathrm{SU}(2))$ and its associated bundle $E = P \times_\rho \mathbb{C}^2$, where we have taken the two-dimensional representation. The gauge potential on a chart $U_i$ is

$$\mathcal{A}_i = \mathcal{A}_{i\mu}\,\mathrm{d}x^\mu = A_{i\mu}{}^\alpha\left(\frac{\sigma_\alpha}{2\mathrm{i}}\right)\mathrm{d}x^\mu \tag{10.64}$$

where $\sigma_\alpha/2\mathrm{i}$ are generators of SU(2), $\sigma_\alpha$ being the Pauli matrices. Let $e_\alpha{}^0$ $(\alpha = 1, 2)$ be basis vectors of $\mathbb{C}^2$ and consider sections

$$e_\alpha(p) \equiv [(\sigma_i(p), e_\alpha{}^0)] \tag{10.65}$$

where $\sigma_i(p)$ defines a canonical trivialization of $P$ over $U_i$. Let $\phi(p) = [(\sigma_i(p), \Phi^\alpha(p)e_\alpha{}^0)]$ be a section of $E$ over $M$. Along a horizontal lift $\tilde{\gamma}(t) = \sigma_i(p)U(t)$, $U(t) \in \mathrm{SU}(2)$, we have

$$\phi(t) = [(\tilde{\gamma}(t), U(t)^{-1}\Phi^\alpha(t)e_\alpha{}^0)]. \tag{10.66}$$

The covariant derivative of $\phi$ along $X = \mathrm{d}/\mathrm{d}t$ is

$$\begin{aligned}\nabla_X \phi &= \left[\left(\tilde{\gamma}(0), U(0)^{-1}\frac{\mathrm{d}\Phi^\alpha(0)}{\mathrm{d}t}e_\alpha{}^0\right)\right] \\ &\quad + [(\tilde{\gamma}(0), U(0)^{-1}\mathcal{A}_i(X)^\alpha{}_\beta\Phi^\beta(0)e_\alpha{}^0)] \\ &= X^\mu\left(\frac{\partial\Phi^\alpha}{\partial x^\mu} + \mathcal{A}_{i\mu}{}^\alpha{}_\beta\Phi^\beta\right)e_\alpha\end{aligned} \tag{10.67}$$

where (10.13b) has been used to obtain the last equality.

*Exercise 10.10.* Let us consider an associated adjoint bundle $E_{\mathfrak{g}} = P \times_{\mathrm{Ad}} \mathfrak{g}$ where the action of $G$ on $\mathfrak{g}$ is the adjoint action $V \to \mathrm{Ad}_g V = g^{-1}Vg$, $V \in \mathfrak{g}$ and $g \in G$. Take a local section $\sigma_i \in \Gamma(U_i, P)$ such that $\tilde{\gamma}(t) = \sigma_i(t)g(t)$. Take a section $s(p) = [(\sigma_i(p), V(p))]$ on $E_{\mathfrak{g}}$, where $V(p) = V^\alpha(p)T_\alpha$, $\{T_\alpha\}$ being the basis of $\mathfrak{g}$. Define the covariant derivative $\mathcal{D}_X s$ by

$$\mathcal{D}_X s \equiv \left[\left(\tilde{\gamma}(0), \frac{\mathrm{d}}{\mathrm{d}t}\{\mathrm{Ad}_{g(t)^{-1}} V(t)\}\Big|_{t=0}\right)\right]. \tag{10.68a}$$

Show that

$$\begin{aligned}\mathcal{D}_X s &= \left[\left(\sigma_i(0), \frac{\mathrm{d}V(t)}{\mathrm{d}t} + [\mathcal{A}_i(X), V(t)]\Big|_{t=0}\right)\right] \\ &= X^\mu \left(\frac{\partial V^\alpha}{\partial x^\mu} + f_{\beta\gamma}{}^\alpha \mathcal{A}_{i\mu}{}^\beta V^\gamma\right)[(\sigma_i(0), T_\alpha)].\end{aligned} \tag{10.68b}$$

### 10.4.3 Curvature rederived

The covariant derivative $\nabla_X s$ defines an operator $\nabla : \Gamma(M, E) \to \Gamma(M, E \otimes \Omega^1(M))$ by (10.50). More generally, the action of $\nabla$ on a vector-valued $p$-form $s \otimes \eta$, $\eta \in \Omega^p(M)$, is defined by

$$\nabla(s \otimes \eta) \equiv (\nabla s) \wedge \eta + s \otimes \mathrm{d}\eta. \tag{10.69}$$

Let $U_i$ be a chart of $M$ and $\sigma_i$ a section of $P$ over $U_i$. We take the canonical local trivialization over $U_i$. We now prove

$$\nabla\nabla e_\alpha = e_\beta \otimes \mathcal{F}_i{}^\beta{}_\alpha \tag{10.70}$$

where $e_\alpha = [(\sigma_i, e_\alpha{}^0)] \in \Gamma(U_i, E)$. In fact, by straightforward computation, we find

$$\begin{aligned}\nabla\nabla e_\alpha &= \nabla(e_\beta \otimes \mathcal{A}_i{}^\beta{}_\alpha) = \nabla e_\beta \wedge \mathcal{A}_i{}^\beta{}_\alpha + e_\beta \otimes \mathrm{d}\mathcal{A}_i{}^\beta{}_\alpha \\ &= e_\beta \otimes (\mathrm{d}\mathcal{A}_i{}^\beta{}_\alpha + \mathcal{A}_i{}^\beta{}_\gamma \wedge \mathcal{A}_i{}^\gamma{}_\alpha) = e_\beta \otimes \mathcal{F}_i{}^\beta{}_\alpha.\end{aligned}$$

*Exercise 10.11.* Let $s(p) = \xi^\alpha(p)e_\alpha(p)$ be a section of $E$. Show that

$$\nabla\nabla s = e_\alpha \otimes \mathcal{F}_i{}^\alpha{}_\beta \xi^\beta. \tag{10.71}$$

### 10.4.4 A connection which preserves the inner product

Let $E \xrightarrow{\pi} M$ be a vector bundle with a positive-definite symmetric inner product whose action is defined at each point $p \in M$ by

$$g_p : \pi^{-1}(p) \otimes \pi^{-1}(p) \to \mathbb{R}. \tag{10.72}$$