Differential Geometry

36

In the last two chapters, we introduced the notions of the principal fiber bundle and its associated bundle. The former made contact with physics by the introduction of a connection—identified as the gauge potential—and its curvature—identified as the gauge field. The latter, the associated bundle with a vector space as its standard fiber, was the convenient setting for particle fields. The concentration of Chap. 35 was on the objects, the particle fields $\boldsymbol{\psi}$, that lived in the vector space and not on the vector space itself. The importance of the vector space comes from the fact that tangent spaces of the base manifold M are vector spaces, and their examination leads to the nature of the base manifold. And that is our aim in this chapter.

36.1 Connections in a Vector Bundle

Let $P(M, G)$ be a principal fiber bundle and $E(M, \mathbb{R}^m, G, P)$ its associated bundle, where G acts on $\mathbb{R}^m$ by a representation of G into $GL(m, \mathbb{R})$. In such a situation, E is called a **vector bundle**. The set of sections $\mathcal{S}(E, M, \mathbb{R}^m)$ has a natural vector space structure with obvious addition of vectors and multiplication by scalars. Furthermore, if $\lambda \in \mathcal{C}^\infty(M)$, we have

vector bundle

$$(\lambda\varphi)(x) = \lambda(x) \cdot \varphi(x), \quad \varphi \in \mathcal{S}\big(E, M, \mathbb{R}^m\big),\ x \in M.$$

If φ is a section in E, and if φ is to have any physical application, we have to know how to calculate its partial (or directional) derivatives. This means being able to define a differentiation process for φ given a vector field $\mathbf{X}$ on M. For a vector field $\mathbf{X}$, let $\gamma_X(t)$ be its integral curve in the neighborhood of $t = 0$. Denote this curve by x_t, so that $\mathbf{X} = \dot{x}_0$. Lift this curve up to w_t, and see how φ changes along this curve. The derivative of φ along w_t is what we are after.

Definition 36.1.1 Let φ be a section of E defined on the curve $\gamma = x_t$ in M. Let $\dot{x}_t$ be the vector tangent to γ at x_t. The **covariant derivative** $\nabla_{\dot{x}_t}\varphi$ of φ in the direction of (or with respect to) $\dot{x}_t$ is given by

covariant derivative

$$\nabla_{\dot{x}_t}\varphi = \lim_{h\to 0}\frac{1}{h}\big[\gamma_t^{t+h}\big(\varphi(x_{t+h})\big) - \varphi(x_t)\big]$$

S. Hassani, *Mathematical Physics*, DOI 10.1007/978-3-319-01195-0_36,

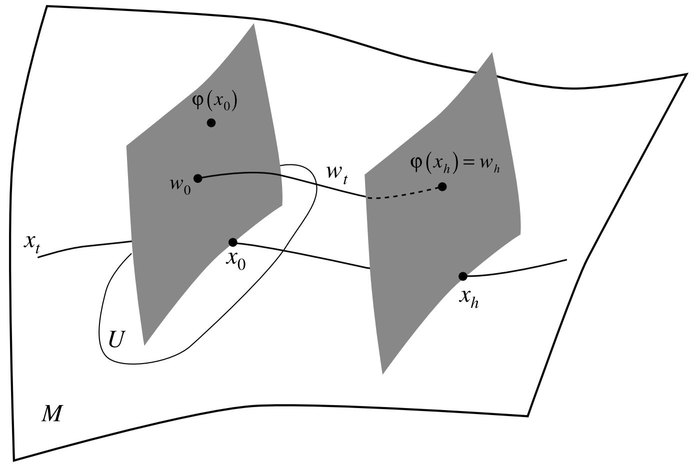

Fig. 36.1 The covariant derivative $\nabla_{\dot{x}_0}\varphi$. The *grey sheets* are $\pi_E^{-1}(x_0)$ and $\pi_E^{-1}(x_h)$, the fibers of E at x_0 and x_h. Find $\varphi(x_h)$. Construct the horizontal lift w_t of x_t starting at $\varphi(x_h)$. Go backwards to the fiber at x_0. Record w_0, and find the difference $\varphi(x_0) - w_0$ and divide by h. As h goes to zero, this ratio gives the covariant derivative of φ with respect to $\dot{x}_0$

where γ_t^{t+h} is the parallel displacement of the fiber $\pi_E^{-1}(x_{t+h})$ along γ to the fiber $\pi_E^{-1}(x_t)$.

Note that $\nabla_{\dot{x}_t}\varphi \in \pi_E^{-1}(x_t)$, and thus defines a section of E along γ.

Remark 36.1.1 Definition 36.1.1 has a number of interesting features which are worth exploring. First, the very notion of derivative involves a subtraction, an operation that does not exist on all mathematical objects such as, for example, manifold. Thus, the fact that fibers of this particular E have a vector-space structure is important. Second, although all fibers are isomorphic as vector spaces, there is no natural isomorphism connecting them. Parallelism gives an isomorphism, but parallelism depends on the notion of a horizontal lift of a curve in the base manifold. Horizontal lift, in turn, depends on the notion of a connection. One interpretation of the word “connection” is that it actually does connect fibers through an induced isomorphism.

Third, any derivative involves an infinitesimal change. Now that we have a curve in the base manifold, it can induce a section-dependent curve $\varphi(x_t)$ in E. A natural directional derivative of the section would be to move along x_t and see how $\varphi(x_t)$ changes. When t changes to $t + h$, the section changes from $\varphi(x_t)$ to $\varphi(x_{t+h})$. But we cannot compute the difference $\varphi(x_{t+h}) - \varphi(x_t)$, because $\varphi(x_{t+h}) \in \pi_E^{-1}(x_{t+h})$ while $\varphi(x_t) \in \pi_E^{-1}(x_t)$, and we don’t know how to subtract two vectors from two different vector spaces. That is why we need to transfer $\varphi(x_{t+h})$ to $\pi_E^{-1}(x_t)$ via the parallelism γ_t^{t+h} (see Fig. 36.1). The word “parallelism” comes about because the horizontal lift of x_t is as parallel to x_t as one can get within the confines of a connection.

Definition 36.1.2 A section φ is **parallel** if $\varphi(x_t)$ is the horizontal lift of x_t. In particular, $\gamma_t^{t+h}(\varphi(x_{t+h})) = \varphi(x_t)$ for all t and h. parallel section

We thus have

Proposition 36.1.3 *A section φ is parallel iff $\nabla_{\dot{x}_t}\varphi = 0$ for all t.*

Furthermore, if we rewrite the defining equation of the covariant derivative as

$$\gamma_t^{t+h}\big(\varphi(x_{t+h})\big) = \varphi(x_t) + h\nabla_{\dot{x}_t}\varphi + O\big(h^2\big),$$

where $O(h^2)$ denotes terms of order h^2 and higher powers of h, then we see that two curves that have the same value and tangent vector at t give the same covariant derivative at t. This means that we can define the covariant derivative in terms of vectors.

Definition 36.1.4 Let $\mathbf{X} \in T_x M$ and φ a section of E defined on a neighborhood of $x \in M$. Let x_t be a curve such that $\mathbf{X} = \dot{x}_0$. Then the **covariant derivative** of φ **in the direction of X** is $\nabla_X\varphi = \nabla_{\dot{x}_0}\varphi$. A section φ is parallel on $U \subset M$ iff $\nabla_X\varphi = 0$ for all $\mathbf{X} \in T_x U$, $x \in U$. covariant derivative in the direction of a vector

It is convenient to have an alternative definition of the covariant derivative of a section in the direction of the vector $\mathbf{X}$ in terms of its horizontal lift in P.

Proposition 36.1.5 *Let φ be defined on $U \subset M$. Associate with φ an $\mathbb{R}^m$-valued function f on $\pi^{-1}(U) \subset P$ by $f(p) = p^{-1}(\varphi(\pi(p)))$ for $p \in \pi^{-1}(U) \subset P$ as in Remark* 34.3.1. *Let $\mathbf{X}^* \in T_p P$ be the horizontal lift of $\mathbf{X} \in T_x M$. Then $\mathbf{X}^* f \in \mathbb{R}^m$ and $p(\mathbf{X}^* f) \in \pi^{-1}(x)$, and*

$$\nabla_X\varphi = p\big(\mathbf{X}^* f\big).$$

Proof Let x_t be the curve with $\dot{x}_0 = \mathbf{X}$. Let p_t be the horizontal lift of x_t such that $\mathbf{X}^* = \dot{p}_0$ and $p_0 = p$. Then we have

$$\mathbf{X}^* f = \lim_{h\to 0}\frac{1}{h}\big[f(p_h) - f(p)\big] = \lim_{h\to 0}\frac{1}{h}\big[p_h^{-1}\big(\varphi(x_h)\big) - p^{-1}\big(\varphi(x)\big)\big]$$

and

$$p\mathbf{X}^* f = \lim_{h\to 0}\frac{1}{h}\big[pp_h^{-1}\big(\varphi(x_h)\big) - \varphi(x)\big].$$

Set $\xi = p_h^{-1}(\varphi(x_h))$, and consider $p_t\xi$, which is a horizontal curve in E. Note that $p_h\xi = \varphi(x_h)$ and $p_0\xi = pp_h^{-1}(\varphi(x_h))$. By the definition of γ_0^h, we have $\gamma_0^h(p_h\xi) = p_0\xi$. Hence, $\gamma_0^h\varphi(x_h) = pp_h^{-1}(\varphi(x_h))$. Substituting this in the above equation yields the result we are after. □

Sometimes it is convenient to write the definition of the covariant derivative in terms of ordinary derivatives, as follows

$$\nabla_{\dot{x}_t}\varphi = \frac{d}{ds}\gamma_t^{t+s}\big(\varphi(x_{t+s})\big)\Big|_{s=0}. \tag{36.1}$$

The covariant derivative satisfies certain important properties which are sometimes used to define it. We collect these properties in the following

Proposition 36.1.6 *Let* $\mathbf{X}, \mathbf{Y} \in T_x M$ *and* φ *and* ψ *be sections of* E *defined in a neighborhood of* x*. Then*

(a) $\nabla_X(\varphi + \psi) = \nabla_X\varphi + \nabla_X\psi$;
(b) $\nabla_{\alpha X}\varphi = \alpha\nabla_X\varphi$, *where* $\alpha \in \mathbb{R}$;
(c) $\nabla_{X+Y}\varphi = \nabla_X\varphi + \nabla_Y\varphi$;
(d) $\nabla_X(f\varphi) = f(x)\cdot\nabla_X\varphi + (\mathbf{X}f)\cdot\varphi(x)$, *where* f *is a real-valued function defined in a neighborhood of* x*.*

Proof (a) follows from the fact that the isomorphism γ_t^{t+h} is linear. (b) follow from the fact that if $\gamma_X(t)$ is the curve whose tangent is $\mathbf{X}$, then $\gamma_X(\alpha t)$ is the curve whose tangent is $\alpha\mathbf{X}$. (c) follows from Proposition 36.1.5 and the fact that $\mathbf{X}^* + \mathbf{Y}^*$ is the lift of $\mathbf{X} + \mathbf{Y}$. For (d), let $\mathbf{X}$ be tangent to x_t with $x_0 = x$ and $\mathbf{X} = \dot{x}_0$. Then use Eq. (36.1):

$$\begin{aligned}
\nabla_X(f\varphi) = \nabla_{\dot{x}_0}(f\varphi) &= \frac{d}{dt}\gamma_0^t\big(f(x_t)\varphi(x_t)\big)\Big|_{t=0} = \frac{d}{dt}f(x_t)\gamma_0^t\big(\varphi(x_t)\big)\Big|_{t=0} \\
&= \frac{d}{dt}f(x)\gamma_0^t\big(\varphi(x_t)\big)\Big|_{t=0} + \frac{d}{dt}f(x_t)\gamma_0^0\big(\varphi(x)\big)\Big|_{t=0} \\
&= f(x)\cdot\nabla_X(\varphi) + (\mathbf{X}f)\cdot\varphi(x)
\end{aligned}$$

where we used $\gamma_0^t(f(x_t)\varphi(x_t)) = f(x_t)\gamma_0^t(\varphi(x_t))$, which is a property of the linearity of γ_0^t and the fact that $f(x_t)$ is a real number. We also used $\gamma_0^0 = \mathrm{id}$, which should be obvious. □

Remark 36.1.2 Proposition 36.1.6 applies to vectors at a point of the manifold M. It can be generalized to vector fields by applying it pointwise. Therefore, the proposition holds for vector fields as well. The minor difference is that α in (b) of the proposition can also be a function on M.

36.2 Linear Connections

From the general vector bundles whose standard fiber was $\mathbb{R}^m$, we now specialize to the bundle of linear frames $L(M)$ examined in Example 34.1.11, among whose associated bundles are tangent bundle (Box 34.1.17) and tensor bundle (Example 34.1.18). We use P for $L(M)$ to avoid cluttering the notations.

Definition 36.2.1 A connection Γ in $L(M) \equiv P$ is called a **linear connection** of M. linear connection

Recall that for any principal fiber bundle and its associate bundle E, each $p \in P$ is an isomorphism of F, the standard fiber of E, with $\pi_E^{-1}(x)$. In fact, p^{-1} is an F-valued map on $\pi_E^{-1}(x)$. In the present case of $L(M)$, p^{-1} is an $\mathbb{R}^n$-valued map on $T_x M$. In addition, there is a natural map $T_p P \to T_x M$, namely π_*. If we combine the two maps, we get a 1-form on $L(M)$:

Definition 36.2.2 The **canonical form** $\boldsymbol{\theta}$ of $L(M) \equiv P$ is the $\mathbb{R}^n$-valued 1-form on P defined by canonical form

$$\boldsymbol{\theta}(\mathbf{X}) = p^{-1}\pi_*(\mathbf{X}) \quad \text{for } \mathbf{X} \in T_p P. \tag{36.2}$$

Proposition 36.2.3 *The canonical form is a tensorial* 1*-form of type* $(\mathrm{id}, \mathbb{R}^n)$, *where* id *is the identity representation of* $GL(n;\mathbb{R})$.

Proof Let $\mathbf{X}$ be any vector at $p \in P$ and $g \in GL(n;\mathbb{R})$. Then $R_{g*}\mathbf{X}$ is a vector at $pg \in P$. Therefore,

$$\begin{aligned}\big(R_g^*\boldsymbol{\theta}\big)(\mathbf{X}) &= \boldsymbol{\theta}(R_{g*}\mathbf{X}) = (pg)^{-1}\big(\pi_*(R_{g*}\mathbf{X})\big)\\ &= g^{-1}p^{-1}\big(\pi_*(\mathbf{X})\big) = g^{-1}\cdot\boldsymbol{\theta}(\mathbf{X}),\end{aligned}$$

where we used $\pi_*(R_{g*}\mathbf{X}) = \pi_*(\mathbf{X})$ which is implied by $\pi(pg) = \pi(p)$. This shows that $\boldsymbol{\theta}$ is pseudotensorial. But if $\mathbf{X}$ is vertical, then π_*—and therefore $\boldsymbol{\theta}$—annihilates it. Hence, $\boldsymbol{\theta}$ is tensorial. □

Definition 36.2.4 Let Γ be a linear connection of M. For each $\xi \in \mathbb{R}^n$ and $p \in P$ define the vector field $\mathbf{B}(\xi)$ in such a way that $(\mathbf{B}(\xi))_p$ is the unique horizontal lift of $p\xi \in T_{\pi(p)}M$. The vector field $\mathbf{B}(\xi)$ so defined is called the **standard horizontal vector field** of Γ corresponding to ξ. standard horizontal vector field of a connection

Proposition 36.2.5 *The standard horizontal vector fields have the following properties*:

(a) *If* $\boldsymbol{\theta}$ *is the canonical form of* P, *then* $\boldsymbol{\theta} \circ \mathbf{B} = \mathrm{id}_{\mathbb{R}^n}$.
(b) $R_{g*}(\mathbf{B}(\xi)) = \mathbf{B}(g^{-1}\xi)$ *for* $g \in GL(n;\mathbb{R})$ *and* $\xi \in \mathbb{R}^n$.
(c) *If* $\xi \neq 0$, *then* $\mathbf{B}(\xi)$ *never vanishes.*

Proof (a) follows directly from the definition of $\boldsymbol{\theta}$. In fact

$$\boldsymbol{\theta}_p\big(\big(\mathbf{B}(\xi)\big)_p\big) = p^{-1}\big(\pi_*\big(\big(\mathbf{B}(\xi)\big)_p\big)\big) = p^{-1}(p\xi) = \xi.$$

(b) If $\mathbf{B}(\xi)$ is a standard horizontal vector field at p, then $R_{g*}(\mathbf{B}(\xi))$ is a standard horizontal vector field at pg. Let $R_{g*}(\mathbf{B}(\xi)) \equiv \mathbf{B}(\xi')$. Then $\pi_*((\mathbf{B}(\xi'))_{pg}) = pg\xi'$. We also have $\pi_*((\mathbf{B}(\xi))_p) = p\xi$. Since $\pi_*(R_{g*}\mathbf{X}) = \pi_*(\mathbf{X})$, we must have $pg\xi' = p\xi$ or $\xi' = g^{-1}\xi$.

(c) Assume that $(\mathbf{B}(\xi))_p = 0$ for some p. Then $p\xi = \pi_*((\mathbf{B}(\xi))_p) = 0$. Multiplying by p^{-1}, we get $\xi = 0$. □

Proposition 36.2.6 *Let A^* be the fundamental vector field corresponding to $\mathsf{A} \in \mathfrak{g}$ and $\mathbf{B}(\xi)$ the standard horizontal vector field corresponding to $\xi \in \mathbb{R}^n$. Then*

$$\left[A^*, \mathbf{B}(\xi)\right] = \mathbf{B}(\mathsf{A}\xi).$$

Proof Recall that the commutator of two vector fields is the Lie derivative of one with respect to the other. Hence, using Definition 28.4.12, noting that the action of G on P is a right action, and using (b) of Proposition 36.2.5, we have

$$\begin{aligned}\left[A^*, \mathbf{B}(\xi)\right] &= \lim_{t\to 0} \frac{1}{t}\left[R_{g_t^{-1}*}\mathbf{B}(\xi) - \mathbf{B}(\xi)\right] \\ &= \lim_{t\to 0} \frac{1}{t}\left[\mathbf{B}(g_t\xi) - \mathbf{B}(\xi)\right] = \lim_{t\to 0} \frac{1}{t}\left[\mathbf{B}\left(e^{t\mathsf{A}}\xi\right) - \mathbf{B}(\xi)\right] \\ &= \lim_{t\to 0} \frac{1}{t}\left[\mathbf{B}\left((1 + t\mathsf{A} + \cdots)\xi\right) - \mathbf{B}(\xi)\right] = \mathbf{B}(\mathsf{A}\xi),\end{aligned}$$

where we used $\exp(t\mathsf{A}) = e^{t\mathsf{A}}$ when the Lie algebra is $\mathfrak{gl}(n;\mathbb{R})$. □

torsion form

Definition 36.2.7 The **torsion form** of a linear connection $\boldsymbol{\omega}$ is defined by

$$\boldsymbol{\Theta} = D^{\omega}\boldsymbol{\theta}. \tag{36.3}$$

Proposition 36.2.3 implies that $\boldsymbol{\Theta} \in \bar{\Lambda}^2(P, \mathbb{R}^n)$, i.e., that the torsion form is a tensorial 2-form.

Theorem 36.2.8 *Let $\boldsymbol{\omega}$, $\boldsymbol{\Theta}$, and $\boldsymbol{\Omega}$ be the connection form, the torsion form, and the curvature form of a linear connection Γ of $L(M)$. Then we have,*

First structure equation: $\boldsymbol{\Theta} = d\boldsymbol{\theta} + \boldsymbol{\omega}\dot{\wedge}\boldsymbol{\theta}$, *or in detail*,

$$\boldsymbol{\Theta}(\mathbf{X}, \mathbf{Y}) = d\boldsymbol{\theta}(\mathbf{X}, \mathbf{Y}) + \boldsymbol{\omega}(\mathbf{X}) \cdot \boldsymbol{\theta}(\mathbf{Y}) - \boldsymbol{\omega}(\mathbf{Y}) \cdot \boldsymbol{\theta}(\mathbf{X})$$

Second structure equation: $\boldsymbol{\Omega} = d\boldsymbol{\omega} + \frac{1}{2}[\boldsymbol{\omega}, \boldsymbol{\omega}]$, *or*

$$\boldsymbol{\Omega}(\mathbf{X}, \mathbf{Y}) = d\boldsymbol{\omega}(\mathbf{X}, \mathbf{Y}) + \frac{1}{2}\left[\boldsymbol{\omega}(\mathbf{X}), \boldsymbol{\omega}(\mathbf{Y})\right],$$

where $\mathbf{X}, \mathbf{Y} \in T_p(L(M))$.

Proof The second structure equation is the result of Theorem 34.3.6. The first structure equation is derived in [Koba 63, pp. 120–121]. □

The equations above which do not act on vector fields are to be interpreted as products of matrices whose entries are forms and the multiplication is through wedge product. We can write the equations above in terms of components. Let $\{\hat{\mathbf{e}}_i\}_{i=1}^n$ be the standard basis of $\mathbb{R}^n$ and $\{\mathbf{E}_i^j\}_{i,j=1}^n$ be a basis of $\mathfrak{gl}(n, \mathbb{R})$. $\mathbf{E}_i^j$ is an $n \times n$ matrix with a 1 at the ith column and jth row and zero everywhere else. In terms of these basis vectors, we can write

$$\begin{aligned} \boldsymbol{\theta} &= \theta^i \hat{\mathbf{e}}_i, & \boldsymbol{\Theta} &= \Theta^i \hat{\mathbf{e}}_i, \\ \boldsymbol{\omega} &= \omega^i_j \mathbf{E}_i^j, & \boldsymbol{\Omega} &= \Omega^i_j \mathbf{E}_i^j, \end{aligned} \tag{36.4}$$

with summation over repeated indices in place. Then the structure equations become

$$\begin{aligned} \Theta^i &= d\theta^i + \omega^i_j \wedge \theta^j, \quad i = 1, 2, \ldots, n, \\ \Omega^i_j &= d\omega^i_j + \omega^i_k \wedge \omega^k_j, \quad i, j = 1, 2, \ldots, n, \end{aligned} \tag{36.5}$$

as the reader can verify (see Problem 36.2). Multiplying both sides of the second equation above by $\mathbf{E}_i^j$, the left-hand side becomes a matrix with elements Ω^i_j. The first terms on the right-hand side becomes the exterior derivative of a matrix whose elements are ω^i_j and the second term on the right-hand side will be simply the matrix product of the latter matrix, where the elements are wedge-multiplied. We summarize this in the following box, which we shall use later:

Box 36.2.9 *Let $\widehat{\boldsymbol{\Omega}}$ be the matrix with elements Ω^i_j and $\widehat{\boldsymbol{\omega}}$ the matrix with elements ω^i_j. Then $\widehat{\boldsymbol{\Omega}} = d\widehat{\boldsymbol{\omega}} + \widehat{\boldsymbol{\omega}} \wedge \widehat{\boldsymbol{\omega}}$, where d operates on the elements of $\widehat{\boldsymbol{\omega}}$ and $\widehat{\boldsymbol{\omega}} \wedge \widehat{\boldsymbol{\omega}}$ is the multiplication of two matrices in which ordinary product is replaced with wedge product.*

Theorem 36.2.10 (Bianchi's identities) *Let $\boldsymbol{\omega}$, $\boldsymbol{\Theta}$, and $\boldsymbol{\Omega}$ be the connection form, the torsion form, and the curvature form of a linear connection Γ of $L(M)$. Then*

First Bianchi identity: $D^\omega \boldsymbol{\Theta} = \boldsymbol{\Omega} \dot{\wedge} \boldsymbol{\theta}$.
Second Bianchi identity: $D^\omega \boldsymbol{\Omega}^\omega = 0$.

Proof The first identity is a special case of Theorem 35.1.4. The second identity was the content of Theorem 34.3.7. □

36.2.1 Covariant Derivative of Tensor Fields

Up to this point, we have concentrated on the differentiation of forms, whose natural differential is D^ω. We also need to differentiate general tensors in the most "natural" way. As discussed earlier, this natural way is the directional derivative introduced in Proposition 36.1.6. However, instead of derivatives with respect to a vector at a point, we generalize to derivatives in the direction of a vector *field* as pointed out in Remark 36.1.2. When the standard fiber is $\mathbb{R}^n$, with $n = \dim M$, the associated bundle is the tangent bundle $T(M)$ whose cross sections are vector fields. We thus restate Proposition 36.1.6 for $L(M)$ with the associated bundle $T(M)$ in the direction of vector fields:

Proposition 36.2.11 *Let* $\mathbf{X}$, $\mathbf{Y}$, *and* $\mathbf{Z}$ *be vector fields on* M. *Then*

(a) $\nabla_X(\mathbf{Y}+\mathbf{Z}) = \nabla_X\mathbf{Y} + \nabla_X\mathbf{Z}$;
(b) $\nabla_{X+Y}\mathbf{Z} = \nabla_X\mathbf{Z} + \nabla_Y\mathbf{Z}$;
(c) $\nabla_{fX}\mathbf{Y} = f\cdot\nabla_X\mathbf{Y}$, *where* $f \in \mathcal{C}^\infty(M)$;
(d) $\nabla_X(f\mathbf{Y}) = f\cdot\nabla_X\mathbf{Y} + (\mathbf{X}f)\cdot\mathbf{Y}$, *where* $f \in \mathcal{C}^\infty(M)$.

Any derivative satisfying (a)–(d) of Proposition 36.2.11 is the covariant derivative with respect to a linear connection.

We defined the covariant derivative in terms of parallel displacements along a path in M and obtained the four equations of Proposition 36.2.11. It turns out that

Theorem 36.2.12 *Any derivative which satisfies the four conditions of Proposition 36.2.11 is the covariant derivative with respect to some linear connection.*

If instead of $\mathbb{R}^n$, we take $\mathcal{T}_s^r(\mathbb{R}^n)$ as the standard basis, the bundle associated with $L(M)$ becomes the tensor bundle $T_s^r(M)$ of type (r,s) over M. Being still a vector bundle, we can define a covariant derivative for it. Now, tensors are products of vector fields and 1-forms, and if we know how the directional derivative acts on vector fields and one forms, we know how it acts on all tensors. Since a 1-form pairs up with a vector field to produce a function, we can also state that if the action of the derivative is known for vector fields and functions, it is known for all tensors. The action of ∇_X on vector fields is given by Proposition 36.2.11. The proposition also includes its action on functions (see Problem 36.3).

Theorem 36.2.13 *Let* $\mathcal{T}(M)$ *be the algebra of tensor fields* (*the vector space of tensor fields together with tensor multiplication as the algebra multiplication*) *on* M. *Then the covariant differentiation has the following properties*:

(a) $\nabla_X : \mathcal{T}(M) \to \mathcal{T}(M)$ *is a type-preserving derivation*;[1]
(b) ∇_X *commutes with contraction*;
(c) $\nabla_X f = \mathbf{X}f$ *for every function* $f \in \mathcal{C}^\infty(M)$ *on* M;
(d) $\nabla_{X+Y} = \nabla_X + \nabla_Y$
(e) $\nabla_{fX}\mathbf{T} = f\cdot\nabla_X\mathbf{T}$ *for all* $f \in \mathcal{C}^\infty(M)$ *and* $\mathbf{T} \in \mathcal{T}(M)$.

A tensor field $\mathbf{T}$ of type (r,s) can be thought of as a multilinear mapping

$$\mathbf{T} : \underbrace{T(M)\times T(M)\times\cdots\times T(M)}_{s \text{ times Cartesian product}} \to \mathcal{T}_0^r(M).$$

The s vector fields from the domain of $\mathbf{T}$ fill all the covariant slots, leaving all the r contravariant slots untouched. With this kind of interpretation, we have the following:

covariant differential

[1]See Definition 3.4.1.

Definition 36.2.14 Given a tensor field $\mathbf{T}$ of type (r, s) the **covariant differential** $\nabla\mathbf{T}$ of $\mathbf{T}$ is a tensor of type $(r, s+1)$ given by

$$(\nabla\mathbf{T})(\mathbf{X}_1, \dots, \mathbf{X}_s; \mathbf{X}) = (\nabla_X\mathbf{T})(\mathbf{X}_1, \dots, \mathbf{X}_s), \qquad \mathbf{X}_i, \mathbf{X} \in T(M).$$

By a procedure similar to that which led to the Lie derivative (28.36) of a p-form, we can obtain the following formula:

$$(\nabla\mathbf{T})(\mathbf{X}_1, \dots, \mathbf{X}_s; \mathbf{X}) = \nabla_X\big(\mathbf{T}(\mathbf{X}_1, \dots, \mathbf{X}_s)\big) - \sum_{i=1}^{s}\mathbf{T}(\mathbf{X}_1, \dots, \nabla_X\mathbf{X}_i, \dots, \mathbf{X}_s), \qquad (36.6)$$

where $\mathbf{T}$ is a tensor field of type (r, s) and $\mathbf{X}_i, \mathbf{X} \in T(M)$.

A tensor field $\mathbf{T}$, as a section of a bundle associated with $L(M)$, is said to be parallel iff $\nabla_X\mathbf{T} = 0$ for all vector fields on M (see Proposition 36.1.3). This leads to

Proposition 36.2.15 *A tensor field* $\mathbf{T}$ *on* M *is parallel iff* $\nabla\mathbf{T} = 0$.

36.2.2 From Forms on P to Tensor Fields on M

We have defined two kinds of covariant differentiation: D^ω, which acts on differential forms on P, and ∇, which acts on the sections of the associated bundle E. In general, there is no natural relation between the two, because the standard fiber $\mathbb{R}^m$ of E has no relation to the structure of P. However, when $P = L(M)$ and the standard fiber is $\mathbb{R}^n$, the fibers $\pi_E^{-1}(x)$ become T_xM, the tangent spaces of the base manifold of the bundle, which are reachable by π_*. Therefore, we expect some kind of a relationship between the two covariant derivatives. In particular, the two quantities defined in terms of D^ω, namely the torsion and curvature forms, should be related to quantities defined in terms of ∇.

The torsion form, being an $\mathbb{R}^n$-valued 2-form on $P = L(M)$, takes two vector fields on P and produces a vector in $\mathbb{R}^n$, the standard fiber of E. Then, through the action of $p \in P$, this vector can be mapped to a vector in T_xM with $x = \pi(p)$ as in Theorem 34.1.14. This process, in conjunction with Remark 34.3.1 allows us to define a map $\mathbf{T} : T(M) \times T(M) \to T(M)$. This map is called **torsion tensor field** or just **torsion**, and is defined as follows: torsion

$$\mathbf{T}(\mathbf{X}, \mathbf{Y}) \equiv p\big(\Theta\big(\mathbf{X}^*, \mathbf{Y}^*\big)\big) \quad \text{for } \mathbf{X}, \mathbf{Y} \in T_xM, \qquad (36.7)$$

where p is any point of $L(M)$ such that $\pi(p) = x$, and $\mathbf{X}^*$ and $\mathbf{Y}^*$ are any two vectors of $L(M)$ such that $\pi_*(\mathbf{X}^*) = \mathbf{X}$ and $\pi_*(\mathbf{Y}^*) = \mathbf{Y}$. Remark 34.3.1 ensures that $\mathbf{T}(\mathbf{X}, \mathbf{Y})$ is independent of p, $\mathbf{X}^*$, and $\mathbf{Y}^*$. Furthermore, $\mathbf{T}(\mathbf{X}, \mathbf{Y}) = -\mathbf{T}(\mathbf{Y}, \mathbf{X})$, and since it maps $T(M) \times T(M)$ to $T(M)$, it is a skew-symmetric tensor field of type $(1, 2)$.

Similarly, The curvature form, being a $\mathfrak{gl}(n, \mathbb{R})$-valued 2-form on $P = L(M)$, takes two vector fields on P and produces a matrix in $\mathfrak{gl}(n, \mathbb{R})$. This matrix can act on a vector in $\mathbb{R}^n$, which can be the inverse map of Theorem 34.1.14 (i.e., the image of a vector in $T_x M$ by p^{-1}), to produce another vector in $\mathbb{R}^n$. Then, through the action of $p \in P$, this vector can be mapped to a vector in $T_x M$ with $x = \pi(p)$ as in Theorem 34.1.14. This process, in conjunction with Remark 34.3.1 allows us to define a map $T(M) \times T(M) \times T(M) \to T(M)$. This map is called **curvature tensor field** or just **curvature**, and is defined as follows:

curvature

$$\mathbf{R}(\mathbf{X}, \mathbf{Y})\mathbf{Z} \equiv p\big(\boldsymbol{\Omega}\big(\mathbf{X}^*, \mathbf{Y}^*\big)\big(p^{-1}\mathbf{Z}\big)\big) \quad \text{for } \mathbf{X}, \mathbf{Y}, \mathbf{Z} \in T_x M. \tag{36.8}$$

It follows that $\mathbf{R}$ is a tensor field of type $(1, 3)$ with the property that $\mathbf{R}(\mathbf{X}, \mathbf{Y}) = -\mathbf{R}(\mathbf{Y}, \mathbf{X})$. Note that $\mathbf{R}(\mathbf{X}, \mathbf{Y})$ is an endomorphism of $T_x M$, and is called the **curvature transformation** of $T_x M$ determined by $\mathbf{X}$ and $\mathbf{Y}$.

curvature transformation

Theorem 36.2.16 *The torsion* $\mathbf{T}$ *and curvature* $\mathbf{R}$ *can be expressed in terms of covariant differentiation as follows*:

$$\mathbf{T}(\mathbf{X}, \mathbf{Y}) = \nabla_X \mathbf{Y} - \nabla_Y \mathbf{X} - [\mathbf{X}, \mathbf{Y}]$$
$$\mathbf{R}(\mathbf{X}, \mathbf{Y})\mathbf{Z} = [\nabla_X, \nabla_Y]\mathbf{Z} - \nabla_{[X,Y]}\mathbf{Z},$$

where $\mathbf{X}$, $\mathbf{Y}$, *and* $\mathbf{Z}$ *are vector fields on* M.

Proof Specialize Proposition 36.1.5 to $L(M)$ and get $(\nabla_X \mathbf{Y})_x = p(\mathbf{X}_p^* f)$ where $\mathbf{X}_p^*$ is the horizontal lift of $\mathbf{X}$ at x. The function f is given by

$$f(p) = p^{-1}(\mathbf{Y}_x) = p^{-1}\big(\pi_*\big(\mathbf{Y}_p^*\big)\big) \equiv \boldsymbol{\theta}\big(\mathbf{Y}_p^*\big), \tag{36.9}$$

where $\mathbf{Y}_p^*$ is the horizontal lift of $\mathbf{Y}$ at x. Thus, we have the identity

$$(\nabla_X \mathbf{Y})_x = p\big(\mathbf{X}_p^*\big(\boldsymbol{\theta}\big(\mathbf{Y}_p^*\big)\big)\big). \tag{36.10}$$

From (36.9), we get $p^{-1}(\nabla_X \mathbf{Y})_x = \boldsymbol{\theta}((\nabla_X \mathbf{Y})_p^*)$. From (36.10), we get $p^{-1}(\nabla_X \mathbf{Y})_x = \mathbf{X}_p^*(\boldsymbol{\theta}(\mathbf{Y}_p^*))$. Therefore, we have another useful identity:

$$\boldsymbol{\theta}\big((\nabla_X \mathbf{Y})^*\big)_p = \mathbf{X}_p^*\big(\boldsymbol{\theta}\big(\mathbf{Y}^*\big)\big). \tag{36.11}$$

Now use the first structure equation of Theorem 36.2.8 to get $\boldsymbol{\Theta}(\mathbf{X}^*, \mathbf{Y}^*) = d\boldsymbol{\theta}(\mathbf{X}^*, \mathbf{Y}^*)$ because $\boldsymbol{\omega}(\mathbf{X}^*) = 0 = \boldsymbol{\omega}(\mathbf{Y}^*)$ for horizontal $\mathbf{X}^*$ and $\mathbf{Y}^*$. We therefore have

$$\begin{aligned}
\mathbf{T}(\mathbf{X}_x, \mathbf{Y}_x) &= p\big(\boldsymbol{\Theta}\big(\mathbf{X}_p^*, \mathbf{Y}_p^*\big)\big) = p\big(d\boldsymbol{\theta}\big(\mathbf{X}_p^*, \mathbf{Y}_p^*\big)\big) \\
&= p\big(\mathbf{X}_p^*\big(\boldsymbol{\theta}\big(\mathbf{Y}^*\big)\big) - \mathbf{Y}_p^*\big(\boldsymbol{\theta}\big(\mathbf{X}^*\big)\big) - \boldsymbol{\theta}\big(\big[\mathbf{X}^*, \mathbf{Y}^*\big]_p\big)\big) \\
&= (\nabla_X \mathbf{Y})_x - (\nabla_Y \mathbf{X})_x - [\mathbf{X}, \mathbf{Y}]_x,
\end{aligned}$$

where we used Theorem 28.5.11 and the fact that $\pi_*([\mathbf{X}^*, \mathbf{Y}^*]) = [\mathbf{X}, \mathbf{Y}]$.

To prove the curvature tensor equation, let $p^{-1}\mathbf{Z} = f(p) = \theta(\mathbf{Z}^*_p)$ by (36.9). Then, we have

$$\begin{aligned}\mathbf{R}(\mathbf{X}_x, \mathbf{Y}_x)\mathbf{Z}_x &= p\big(\boldsymbol{\Omega}\big(\mathbf{X}^*_p, \mathbf{Y}^*_p\big)\big(f(p)\big)\big) = p\big(-\boldsymbol{\omega}\big(\big[\mathbf{X}^*, \mathbf{Y}^*\big]_p\big)\big(f(p)\big)\big) \\ &= p\big(-\boldsymbol{\omega}\big(v\big[\mathbf{X}^*, \mathbf{Y}^*\big]_p\big)\big(f(p)\big)\big) = p\big(-A\big(f(p)\big)\big), \end{aligned} \tag{36.12}$$

where we used Eq. (34.12) and the fact that $\boldsymbol{\omega}$ annihilates the horizontal component of the bracket (v stands for the vertical component). In the last step, we used (a) of Definition 34.2.1 and denoted by A the element of the Lie algebra that gives rise to $A^*_p \equiv v[\mathbf{X}^*, \mathbf{Y}^*]_p$. Now we note that

$$\begin{aligned}-A\big(f(p)\big) &= \frac{d}{dt}\exp(-At)f(p)\bigg|_{t=0} = \frac{d}{dt}f\big(p\exp(At)\big)\bigg|_{t=0} = A^*_p f \\ &= v\big[\mathbf{X}^*, \mathbf{Y}^*\big]_p f = \mathbf{X}^*\big(\mathbf{Y}^*_p f\big) - \mathbf{Y}^*\big(\mathbf{X}^*_p f\big) - h\big[\mathbf{X}^*, \mathbf{Y}^*\big]_p f \\ &= \mathbf{X}^*\big(\mathbf{Y}^*_p\big(\boldsymbol{\theta}\big(\mathbf{Z}^*\big)\big)\big) - \mathbf{Y}^*\big(\mathbf{X}^*_p\big(\boldsymbol{\theta}\big(\mathbf{Z}^*\big)\big)\big) - h\big[\mathbf{X}^*, \mathbf{Y}^*\big]_p\big(\boldsymbol{\theta}\big(\mathbf{Z}^*\big)\big).\end{aligned}$$

It now follows that

$$\begin{aligned}&\mathbf{R}(\mathbf{X}_x, \mathbf{Y}_x)\mathbf{Z}_x \\ &\quad= p\big(\mathbf{X}^*\big(\mathbf{Y}^*_p\big(\boldsymbol{\theta}\big(\mathbf{Z}^*\big)\big)\big) - \mathbf{Y}^*\big(\mathbf{X}^*_p\big(\boldsymbol{\theta}\big(\mathbf{Z}^*\big)\big)\big) - h\big[\mathbf{X}^*, \mathbf{Y}^*\big]_p\big(\boldsymbol{\theta}\big(\mathbf{Z}^*\big)\big)\big) \\ &\quad= p\big(\mathbf{X}^*_p\big(\boldsymbol{\theta}\big((\nabla_Y\mathbf{Z})^*\big)\big) - \mathbf{Y}^*_p\big(\boldsymbol{\theta}\big((\nabla_X\mathbf{Z})^*\big)\big) - h\big[\mathbf{X}^*, \mathbf{Y}^*\big]_p\big(\boldsymbol{\theta}\big(\mathbf{Z}^*\big)\big)\big) \\ &\quad= \nabla_X\nabla_Y\mathbf{Z} - \nabla_Y\nabla_X\mathbf{Z} - \nabla_{[X,Y]}\mathbf{Z},\end{aligned}$$

where we used (36.11) to go from the first line to the second. □

We also want to express the Bianchi's identities in terms of tensor fields. We shall confine ourselves to the case where the torsion form is zero. In most physical applications this is indeed the case. So the first identity of Theorem 36.2.10 becomes $0 = \boldsymbol{\Omega}\dot{\wedge}\boldsymbol{\theta}$. Now let $\mathbf{X}^*$, $\mathbf{Y}^*$, and $\mathbf{Z}^*$ be the lifts of $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$. Then, it is easily shown that

$$0 = (\boldsymbol{\Omega}\dot{\wedge}\boldsymbol{\theta})\big(\mathbf{X}^*, \mathbf{Y}^*, \mathbf{Z}^*\big) = \mathrm{Cyc}\big(\boldsymbol{\Omega}\big(\mathbf{X}^*, \mathbf{Y}^*\big)\boldsymbol{\theta}\big(\mathbf{Z}^*\big)\big),$$

where Cyc means taking the sum of the cyclic permutations of the expression in parentheses. From Eq. (36.12) and the discussion before it, we also see that

$$\mathbf{R}(\mathbf{X}, \mathbf{Y})\mathbf{Z} = p\big(\boldsymbol{\Omega}\big(\mathbf{X}^*, \mathbf{Y}^*\big)\big(\boldsymbol{\theta}\big(\mathbf{Z}^*\big)\big)\big). \tag{36.13}$$

Putting the last two equations together, we obtain

$$\mathrm{Cyc}\big(\mathbf{R}(\mathbf{X}, \mathbf{Y})\mathbf{Z}\big) \equiv \mathbf{R}(\mathbf{X}, \mathbf{Y})\mathbf{Z} + \mathbf{R}(\mathbf{Z}, \mathbf{X})\mathbf{Y} + \mathbf{R}(\mathbf{Y}, \mathbf{Z})\mathbf{X} = 0. \tag{36.14}$$

The proof of the second Bianchi's identity is outlined in Problem 36.5.

Theorem 36.2.17 *Let* $\mathbf{R}$ *be the curvature of a linear connection of* M *whose torsion is zero. Then for* $\mathbf{X}$, $\mathbf{Y}$ *and* $\mathbf{Z}$ *in* $T(M)$, *we have*

Bianchi's 1st identity: $\mathrm{Cyc}[\mathbf{R}(\mathbf{X},\mathbf{Y})\mathbf{Z}] = 0$;
Bianchi's 2nd identity: $\mathrm{Cyc}[\nabla_X \mathbf{R}(\mathbf{Y},\mathbf{Z})] = 0$.

For the sake of completeness, we give the Bianchi's identities for the case where torsion is not zero and refer the reader to [Koba 63, pp. 135–136] for a proof.

Theorem 36.2.18 *Let* $\mathbf{R}$ *and* $\mathbf{T}$ *be, respectively, the curvature and torsion of a linear connection of* M*. Then for* $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z}$ *in* $T(M)$*, we have*

Bianchi's 1st identity: $\mathrm{Cyc}[\mathbf{R}(\mathbf{X},\mathbf{Y})\mathbf{Z}] = \mathrm{Cyc}[\mathbf{T}(\mathbf{T}(\mathbf{X},\mathbf{Y}),\mathbf{Z}) + (\nabla_X\mathbf{T})(\mathbf{Y},\mathbf{Z})]$;
Bianchi's 2nd identity: $\mathrm{Cyc}[\nabla_X\mathbf{R}(\mathbf{Y},\mathbf{Z}) + \mathbf{R}(\mathbf{T}(\mathbf{X},\mathbf{Y}),\mathbf{Z})] = 0$.

36.2.3 Component Expressions

Any application of connections and curvatures requires expressing them in local coordinates. So, it is useful to have the components and identities in which they participate in terms of local coordinates. For a linear bundle, the isomorphism $\pi^{-1}(U) \cong U \times GL(n,\mathbb{R})$ suggests a local coordinate system of the form (x_i, X^j_k), where $(x_1,\dots,x_n)$ is a coordinate system for $U \subset M$ and X^j_k are the elements of a nonsingular $n \times n$ matrix.

Let's start with the canonical 1-form $\boldsymbol{\theta} = \sum_{i=1}^n \theta^i \hat{\mathbf{e}}_i$, with $\hat{\mathbf{e}}_i$ the standard basis of $\mathbb{R}^n$. The most general coordinate expression for θ^i is

$$\theta^i = a^i_j dx^j + b^{ij}_k dX^k_j, \quad \text{therefore} \quad \boldsymbol{\theta} = \big(a^i_j dx^j + b^{ij}_k dX^k_j\big)\hat{\mathbf{e}}_i.$$

By definition, $\boldsymbol{\theta} = p^{-1} \circ \pi_*$, and the presence of π_* annihilates any vertical component of a vector field. This means that $b^{ij}_k = 0$. Now note that as a map $\mathbb{R}^n \to T_xM$, p, whose local coordinates are (x^i, X^j_k), sends $\hat{\mathbf{e}}_k$ to $X^j_k\partial_j$, and p^{-1} sends ∂_k to $Y^i_k\hat{\mathbf{e}}_i$, where Y^i_k is the inverse of X^i_j. Hence, $\boldsymbol{\theta}(\partial_k) = Y^i_k\hat{\mathbf{e}}_i$. So, we have

$$Y^i_k\hat{\mathbf{e}}_i = \boldsymbol{\theta}(\partial_k) = \big(a^i_j dx^j\hat{\mathbf{e}}_i\big)(\partial_k) = a^i_j dx^j(\partial_k)\hat{\mathbf{e}}_i = a^i_j\delta^j_k\hat{\mathbf{e}}_i = a^i_k\hat{\mathbf{e}}_i.$$

Thus, we have our first result:

$$\theta^i = Y^i_k\, dx^k, \quad Y^i_k = \big(X^{-1}\big)^i_k. \tag{36.15}$$

Next, we consider the connection 1-form, $\boldsymbol{\omega} = \omega^i_j\mathbf{E}^j_i$, where ω^i_j are real-valued 1-forms and $\{\mathbf{E}^j_i\}$ form a basis of $\mathfrak{gl}(n,\mathbb{R})$. Write

$$\omega^i_j = a^i_k dX^k_j + b^i_{jk}dx^k$$

and let it act on a fundamental vector field $A_p^* = \frac{d}{dt}(p \exp At)|_{t=0}$. If p is represented by X_β^α, then (with the notation $\partial_\alpha^\beta \equiv \partial/\partial X_\beta^\alpha$)

$$A_p^* = X_\beta^\alpha A_\gamma^\beta \frac{\partial}{\partial X_\gamma^\alpha} = X_\beta^\alpha A_\gamma^\beta \partial_\alpha^\gamma$$

and when ω_j^i acts on A_p^*, it should give A_j^i. So,

$$\begin{aligned} A_j^i = \omega_j^i\big(A_p^*\big) &= a_k^i dX_j^k\big(X_\beta^\alpha A_\gamma^\beta \partial_\alpha^\gamma\big) + b_{jk}^i dx^k\big(X_\beta^\alpha A_\gamma^\beta \partial_\alpha^\gamma\big) \\ &= a_k^i X_\beta^\alpha A_\gamma^\beta dX_j^k\big(\partial_\alpha^\gamma\big) = a_k^i X_\beta^\alpha A_\gamma^\beta \delta_\alpha^k \delta_j^\gamma = a_k^i X_\beta^k A_j^\beta. \end{aligned}$$

For the equality to hold for all A_j^i, we must have $a_k^i X_\beta^k = \delta_\beta^i$, i.e., that $a_k^i = Y_k^i$, where Y_k^i is the inverse of X_k^i. So, we write

$$\omega_j^i = Y_k^i dX_j^k + b_{jk}^i dx^k. \tag{36.16}$$

To get more insight into the composition of b_{jk}^i, let us try to find $\nabla_{\partial_j}\partial_i$ using (36.10). To this end, let $\mathbf{X}_j^*$ be the horizontal lift of ∂_j. The general form of $\mathbf{X}_j^*$ is

$$\mathbf{X}_j^* = \lambda_j^i \partial_i + \Lambda_{j\gamma}^\beta \partial_\beta^\gamma.$$

Since $\pi_*(\mathbf{X}_j^*) = \partial_j$ and $\pi_*(\partial_\beta^\gamma) = 0$, we get $\lambda_j^i = \delta_j^i$. Since $\boldsymbol{\omega}(\mathbf{X}_j^*) = 0 = \omega_m^i(\mathbf{X}_j^*)$, we get

$$\begin{aligned} 0 = \omega_m^i\big(\mathbf{X}_j^*\big) &= Y_k^i dX_m^k\big(\Lambda_{j\gamma}^\beta \partial_\beta^\gamma\big) + b_{mk}^i dx^k(\partial_j) \\ &= Y_k^i \Lambda_{j\gamma}^\beta \delta_\beta^k \delta_m^\gamma + b_{mk}^i \delta_j^k = Y_k^i \Lambda_{jm}^k + b_{mj}^i. \end{aligned}$$

Multiply both sides by X_i^α (and sum over i, of course) to get

$$0 = X_i^\alpha Y_k^i \Lambda_{jm}^k + b_{mj}^i X_i^\alpha = \Lambda_{jm}^\alpha + b_{mj}^i X_i^\alpha \quad \text{or} \quad \Lambda_{jm}^\alpha = -b_{mj}^i X_i^\alpha$$

and

$$\mathbf{X}_j^* = \partial_j - b_{mj}^k X_k^\alpha \partial_\alpha^m. \tag{36.17}$$

With a similar expression for $\mathbf{X}_i^*$. To use (36.10), we need $\boldsymbol{\theta}(\mathbf{X}_i^*)$, which we can obtain using Eq. (36.15), noting that $\boldsymbol{\theta}$ acts only on the horizontal component of $\mathbf{X}_i^*$:

$$\boldsymbol{\theta}\big(\mathbf{X}_i^*\big) = Y_k^\beta\, dx^k(\partial_i)\hat{\mathbf{e}}_\beta = Y_k^\beta \delta_i^k \hat{\mathbf{e}}_\beta = Y_i^\beta \hat{\mathbf{e}}_\beta.$$

As we apply (36.17) to this expression, we must keep in mind that Y_i^β, being the inverse of X_i^β, is independent of x^j. Therefore, only the second term of (36.17) acts on Y_i^β. Then

$$\begin{aligned} \mathbf{X}_j^*\big(\boldsymbol{\theta}\big(\mathbf{X}_i^*\big)\big) &= -b_{mj}^k X_k^\alpha \partial_\alpha^m \big(Y_i^\beta \hat{\mathbf{e}}_\beta\big) = -b_{mj}^k X_k^\alpha \big(\partial_\alpha^m Y_i^\beta\big)\hat{\mathbf{e}}_\beta \\ &= b_{mj}^k X_k^\alpha \big(Y_i^m Y_\alpha^\beta\big)\hat{\mathbf{e}}_\beta = b_{mj}^k Y_i^m \hat{\mathbf{e}}_k. \end{aligned}$$

Finally, we apply p on this and use $p\hat{\mathbf{e}}_k = X^l_k \partial_l$ to obtain

$$\nabla_{\partial_j}\partial_i = b^k_{mj} Y^m_i X^l_k \partial_l \equiv \Gamma^l_{ji}\partial_l, \tag{36.18}$$

where in the last step, we introduced the **Riemann-Christoffel symbols**.[2] These symbols make sense, because $\nabla_{\partial_j}\partial_i$ is a vector field, which could be expanded in terms of the basis $\{\partial_l\}$. The Riemann-Christoffel symbols are simply the coefficients of the expansion. In terms of these symbols, b^i_{jk} can be written as $b^i_{jk} = \Gamma^l_{km} Y^i_l X^m_j$ and Eq. (36.16) can be expressed as

Riemann-Christoffel symbols

$$\omega^i_j = Y^i_k\left(dX^k_j + \Gamma^k_{lm} X^m_j dx^l\right). \tag{36.19}$$

The Riemann-Christoffel symbols could also be obtained from the connection form. From the 1-form ω on P, we define a local $\mathfrak{gl}(n,\mathbb{R})$-valued 1-form ω_U on M as follows. At each point x on M let σ be the section sending x to the linear frame $(\partial_1, \dots, \partial_n)$. A general section sends x to $(X^1_i\partial_1, \dots, X^n_i\partial_n)$ or equivalently, it sends x to the point in P with coordinates (x^i, X^j_k). The particular section we are considering sends x to the point in P with coordinates (x^i, δ^j_k). Moreover, we define $\omega_U = \sigma^*\omega$. Then ω_U is obtained from Eq. (36.19) by setting $X^m_j = \delta^m_j$, $Y^i_k = \delta^i_k$, and $dX^k_j = 0$. Therefore,

$$(\omega_U)^i_j = \Gamma^i_{lj} dx^l. \tag{36.20}$$

We often omit the subscript U when there is no risk of confusion.

Historical Notes

Elwin Bruno Christoffel
1829–1900

Elwin Bruno Christoffel (1829–1900) came from a family in the cloth trade. He attended an elementary school in Montjoie (which was renamed Monschau in 1918) but then spent a number of years being tutored at home in languages, mathematics, and classics. He attended secondary schools from 1844 until 1849. At first he studied at the Jesuit gymnasium in Cologne but moved to the Friedrich-Wilhelms Gymnasium in the same town for at least the three final years of his school education. He was awarded the final school certificate with a distinction in 1849. The next year he went to the University of Berlin and studied under a number of distinguished mathematicians, including Dirichlet.

After one year of military service in the Guards Artillery Brigade, he returned to Berlin to study for his doctorate, which was awarded in 1856 with a dissertation on the motion of electricity in homogeneous bodies. His examiners included mathematicians and physicists, Kummer being one of the mathematics examiners.

At this point Christoffel spent three years outside the academic world. He returned to Montjoie, where his mother was in poor health, but read widely from the works of Dirichlet, Riemann, and Cauchy. It has been suggested that this period of academic isolation had a major effect on his personality and on his independent approach towards mathematics. It was during this time that he published his first two papers on numerical integration, in 1858, in which he generalized Gauss's method of quadrature and expressed the polynomials that are involved as a determinant. This is now called Christoffel's theorem.

In 1859 Christoffel took the qualifying examination to become a university teacher and was appointed a lecturer at the University of Berlin. Four years later, he was appointed to a chair at the Polytechnicum in Zurich, filling the post left vacant when Dedekind went

[2] The reader is cautioned about the order of the lower indices, as it is different in different books.

to Brunswick. Christoffel was to have a huge influence on mathematics at the Polytechnicum, setting up an institute for mathematics and the natural sciences there.
In 1868 Christoffel was offered the chair of mathematics at the Gewerbsakademie in Berlin, which is now the University of Technology of Berlin. However, after three years at the Gewerbsakademie in Berlin, Christoffel moved to the University of Strasbourg as the chair of mathematics, a post he held until he was forced to retire due to ill health in 1892.
Some of Christoffel's early work was on conformal mappings of a simply connected region bounded by polygons onto a circle. He also wrote important papers that contributed to the development of the tensor calculus of Gregorio Ricci-Curbastro and Tullio Levi-Civita. The Christoffel symbols that he introduced are fundamental in the study of tensor analysis. The Christoffel reduction theorem, so named by Klein, solves the local equivalence problem for two quadratic differential forms. The procedure Christoffel employed in his solution of the equivalence problem is what Ricci later called *covariant differentiation*; Christoffel also used the latter concept to define the basic Riemann–Christoffel *curvature tensor*. His approach allowed Ricci and Levi–Civita to develop a coordinate-free differential calculus which Einstein, with the help of Grossmann, turned into the tensor analysis, the mathematical foundation of general relativity.

The Riemann-Christoffel symbols are functions of local coordinates. A change of coordinates transforms the symbols according to a rule that can be easily worked out. In fact, if $\bar{x}^\alpha$ is the new coordinates and $\overline{\Gamma}^\alpha_{\beta\gamma}$ is the symbols in the new coordinate system, then $\nabla_{\bar{\partial}_\beta}\bar{\partial}_\gamma = \overline{\Gamma}^\alpha_{\beta\gamma}\bar{\partial}_\alpha$. To find $\overline{\Gamma}^\alpha_{\beta\gamma}$ in terms of the old symbols, substitute $(\bar{\partial}_\alpha x^i)\partial_i$ for $\bar{\partial}_\alpha$, etc.,

$$\nabla_{\frac{\partial x^i}{\partial \bar{x}^\beta}\partial_i}\left(\frac{\partial x^j}{\partial \bar{x}^\gamma}\partial_j\right) = \overline{\Gamma}^\alpha_{\beta\gamma}\frac{\partial x^k}{\partial \bar{x}^\alpha}\partial_k,$$

then use Proposition 36.2.11 to expand the left-hand side. After some simple manipulation, which we leave for the reader, we obtain

$$\overline{\Gamma}^\alpha_{\beta\gamma} = \Gamma^i_{jk}\frac{\partial x^j}{\partial \bar{x}^\beta}\frac{\partial x^k}{\partial \bar{x}^\gamma}\frac{\partial \bar{x}^\alpha}{\partial x^i} + \frac{\partial^2 x^i}{\partial \bar{x}^\beta \partial \bar{x}^\gamma}\frac{\partial \bar{x}^\alpha}{\partial x^i} \tag{36.21}$$

Because of the presence of the second term, Riemann-Christoffel symbols are *not* components of a tensor.

From the definition of the Riemann-Christoffel symbols, the components of a connection form, we deduced their transformation properties. It turns out that

Theorem 36.2.19 *A set of functions Γ^i_{jk}, which transform according to Eq.* (36.21) *under a change of coordinates, define a unique connection whose components with respect to the coordinates $\{x^i\}$ are Γ^i_{jk}. Furthermore, the connection form $\boldsymbol{\omega} = \omega^i_j \mathbf{E}^j_i$ is given in terms of the local coordinate system by*

$$\omega^i_j = Y^i_k\left(dX^k_j + \Gamma^k_{lm}X^m_j dx^l\right)$$

Proof See [Koba 63, pp. 142–143]. □

Define the components of torsion and curvature tensors by

$$\mathbf{T}(\partial_i, \partial_j) = T^k_{ij}\partial_k, \qquad \mathbf{R}(\partial_i, \partial_j)\partial_k = R^l_{kij}\partial_l. \tag{36.22}$$

Then, using Theorem 36.2.16 one can easily obtain the following

Box 36.2.20 *The components of torsion and curvature tensors are given in terms of the Christoffel symbols*:

$$\begin{aligned} T^k_{ij} &= \Gamma^k_{ij} - \Gamma^k_{ji} \\ R^i_{jkl} &= \partial_k \Gamma^i_{lj} - \partial_l \Gamma^i_{kj} + \Gamma^m_{lj}\Gamma^i_{km} - \Gamma^m_{kj}\Gamma^i_{lm}. \end{aligned} \tag{36.23}$$

In particular, $\Gamma^k_{ji} = \Gamma^k_{ij}$ *if the torsion tensor vanishes.*

Equation (36.20) pulls down the connection 1-form from P to M. One can pull down the curvature 2-form as well. Then Eq. (36.5) gives

$$(\Omega_U)^i_j = d(\omega_U)^i_j + (\omega_U)^i_k \wedge (\omega_U)^k_j. \tag{36.24}$$

As indicated above, we often omit the subscript U. Problem 36.11 shows that

$$\Omega^i_j = \frac{1}{2} R^i_{jkl} dx^k \wedge dx^l \tag{36.25}$$

where R^i_{jkl} is as given in Eq. (36.23).

36.2.4 General Basis

The coordinate expressions derived above express the components of forms and fields in a coordinate basis. We need not confine ourselves to coordinate bases. In fact, they are not always the most convenient bases to use. We can work in a basis $\{\mathbf{e}_i\}$ and its dual $\{\boldsymbol{\epsilon}^j\}$. Then the Riemann-Christofel symbols are defined as

$$\nabla_{\mathbf{e}_i} \mathbf{e}_j = \mathbf{e}_k \Gamma^k{}_{ij}. \tag{36.26}$$

It has to be emphasized that, even in the absence of torsion, in a general basis, $\Gamma^k{}_{ij} \neq \Gamma^k{}_{ji}$. Only in a coordinate basis does this symmetry hold.

Example 36.2.21 Consider two bases $\{\mathbf{e}_i\}$ and $\{\mathbf{e}_{i'}\}$. Write the primed basis in terms of the other: $\mathbf{e}_{i'} = R^j{}_{i'}\mathbf{e}_j$. Then

$$\begin{aligned} \mathbf{e}_{k'}\Gamma^{k'}{}_{i'j'} &\equiv \nabla_{\mathbf{e}_{i'}}\mathbf{e}_{j'} = \nabla_{R^l{}_{i'}\mathbf{e}_l}\left(R^j{}_{j'}\mathbf{e}_j\right) = R^l{}_{i'}\nabla_{\mathbf{e}_l}\left(R^j{}_{j'}\mathbf{e}_j\right) \\ &= R^l{}_{i'}\left\{R^j{}_{j'}\nabla_{\mathbf{e}_l}(\mathbf{e}_j) + \nabla_{\mathbf{e}_l}\left(R^j{}_{j'}\right)\mathbf{e}_j\right\} \\ &= R^l{}_{i'}R^j{}_{j'}\mathbf{e}_m\Gamma^m{}_{lj} + R^l{}_{i'}\underbrace{\mathbf{e}_l\left(R^m{}_{j'}\right)}_{\equiv R^m{}_{j',l}}\mathbf{e}_m. \end{aligned}$$

Connection coefficients are not tensors!

Writing $\mathbf{e}_{k'} = R^m{}_{k'}\mathbf{e}_m$ on the LHS, equating the components on both sides, and multiplying both sides by the inverse of the transformation matrix R, we obtain

$$\Gamma^{k'}{}_{i'j'} = \underbrace{R^{k'}{}_m R^l{}_{i'} R^j{}_{j'} \Gamma^m{}_{lj}}_{\text{how a (1, 2)-tensor transforms}} + \underbrace{R^{k'}{}_m R^l{}_{i'} R^m{}_{j',l}}_{\text{nontensorial term}}, \tag{36.27}$$

where $R^{k'}{}_m \equiv (\mathsf{R}^{-1})_{k'm}$. Equation (36.27) shows once again that the connection coefficients *are not tensors*. Equation (36.21) is a special case of (36.27).

Applying $\nabla_{\mathbf{e}_i}$ to both sides of $\delta^j_m = \langle \mathbf{e}_m, \boldsymbol{\epsilon}^j \rangle$, we obtain

$$\nabla_{\mathbf{e}_i} \boldsymbol{\epsilon}^j = -\Gamma^j{}_{ik} \boldsymbol{\epsilon}^k. \tag{36.28}$$

Since an arbitrary tensor of a given kind can be expressed as a linear combination of tensor product of vectors and 1-forms, Eqs. (36.26) and (36.28), plus the assumed derivation property of $\nabla_{\mathbf{u}}$, is sufficient to uniquely define the action of $\nabla_{\mathbf{u}}$ on any tensor and for any vector field $\mathbf{u}$.

gradient operator for tensors

Let $\mathbf{T}$ be a tensor of type (r, s). The covariant differential of Definition 36.2.14, $\nabla : T^r_s(M) \to T^r_{s+1}(M)$, which is sometimes called the generalized **gradient operator**, when acting on $\mathbf{T}$, adds an extra lower index. In what follows, we want to find the components of $\nabla \mathbf{T}$ in terms of the components of $\mathbf{T}$. If

$$\mathbf{T} = T^{i_1 \dots i_r}_{j_1 \dots j_s} \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_r} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_s},$$

then, following the customary practice of putting the extra lower index after a semicolon, we write

significance of semicolon in components of the covariant derivative of a tensor

$$\nabla \mathbf{T} \equiv T^{i_1 \dots i_r}_{j_1 \dots j_s;k} \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_r} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_s} \otimes \boldsymbol{\epsilon}^k, \tag{36.29}$$

and, with $\mathbf{u} = u^k \mathbf{e}_k$,

$$\nabla_{\mathbf{u}} \mathbf{T} = T^{i_1 \dots i_r}_{j_1 \dots j_s;k} u^k \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_r} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_s}. \tag{36.30}$$

Using these relations, we can calculate the components of the covariant derivative of a general tensor. It is clear that if we use $\mathbf{e}_k$ instead of $\mathbf{u}$, we obtain the kth component of the covariant derivative. So, on the one hand, we have

$$\nabla_{\mathbf{e}_k} \mathbf{T} = T^{i_1 \dots i_r}_{j_1 \dots j_s;k} \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_r} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_s}, \tag{36.31}$$

and on the other hand,

$$\begin{aligned} \nabla_{\mathbf{e}_k} \mathbf{T} &= \nabla_{\mathbf{e}_k} \big(T^{i_1 \dots i_r}_{j_1 \dots j_s} \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_r} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_s} \big) \\ &= T^{i_1 \dots i_r}_{j_1 \dots j_s,k} \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_r} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_s} \\ &\quad + T^{i_1 \dots i_r}_{j_1 \dots j_s} \sum_{m=1}^{r} \mathbf{e}_{i_1} \otimes \cdots \otimes \underbrace{\nabla_{\mathbf{e}_k} \mathbf{e}_{i_m}}_{\mathbf{e}_n \Gamma^n{}_{ki_m}} \otimes \dots \mathbf{e}_{i_r} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_s} \end{aligned}$$

$$+ T^{i_1\dots i_r}_{j_1\dots j_s} \sum_{m=1}^{s} \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_r} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \underbrace{\nabla_{\mathbf{e}_k} \boldsymbol{\epsilon}^{j_m}}_{\substack{-\Gamma^{j_m}_{kn}\boldsymbol{\epsilon}^n \\ \text{by (36.28)}}} \cdots \otimes \boldsymbol{\epsilon}^{j_s}, \tag{36.32}$$

where $T^{i_1\dots i_r}_{j_1\dots j_s,k} \equiv \mathbf{e}_k(T^{i_1\dots i_r}_{j_1\dots j_s})$. Equating the components of Eqs. (36.31) and (36.32) yields

components of the covariant derivative of a tensor

$$T^{i_1\dots i_r}_{j_1\dots j_s;k} = T^{i_1\dots i_r}_{j_1\dots j_s,k} + \sum_{m=1}^{r} T^{i_1\dots i_{m-1} n i_{m+1}\dots i_r}_{j_1\dots j_{m-1} j_m j_{m+1}\dots j_s} \Gamma^{i_m}_{kn} - \sum_{m=1}^{s} T^{i_1\dots i_{m-1} i_m i_{m+1}\dots i_r}_{j_1\dots j_{m-1} n j_{m+1}\dots j_s} \Gamma^{n}{}_{kj_m}, \tag{36.33}$$

where only the sum over the subindex m has been explicitly displayed; the (hidden) sum over repeated indices is, as always, understood. Equation (36.33) is (36.6) written in terms of components.

If $\mathbf{u}$ happens to be tangent to a curve $t \mapsto \gamma(t)$, Eq. (36.30) is written as

$$\nabla_{\mathbf{u}} \mathbf{T} = \frac{DT^{i_1\dots i_r}_{j_1\dots j_s}}{dt} \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_r} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_s}, \tag{36.34}$$

where $DT^{i_1\dots i_r}_{j_1\dots j_s}/dt \equiv T^{i_1\dots i_r}_{j_1\dots j_s;k} u^k$. In a coordinate frame, with $u^i = \dot{x}^i = dx^i/dt$, Eqs. (36.30) and (36.33) give

$$\frac{DT^{i_1\dots i_r}_{j_1\dots j_s}}{dt} = T^{i_1\dots i_r}_{j_1\dots j_s;k} \frac{dx^k}{dt} = \frac{dT^{i_1\dots i_r}_{j_1\dots j_s}}{dt} + \sum_{m=1}^{r} T^{i_1\dots i_{m-1} n i_{m+1}\dots i_r}_{j_1\dots j_{m-1} j_m j_{m+1}\dots j_s} \Gamma^{i_m}_{kn} \frac{dx^k}{dt} - \sum_{m=1}^{s} T^{i_1\dots i_{m-1} i_m i_{m+1}\dots i_r}_{j_1\dots j_{m-1} n j_{m+1}\dots j_s} \Gamma^{n}{}_{kj_m} \frac{dx^k}{dt} \tag{36.35}$$

For the case of a vector (36.35) becomes

$$\frac{Dv^k}{dt} = \frac{dv^k}{dt} + v^j \Gamma^k{}_{ij} \frac{dx^i}{dt}. \tag{36.36}$$

This is an important equation, to which we shall return shortly.

With the generalized gradient operator defined, we can construct the divergence of a tensor just as in vector analysis. Given a vector, the divergence operator $\nabla\cdot$ acts on it and gives a scalar, or, in the language of tensor analysis, it lowers the upper indices by 1. This takes place by differentiating components and contracting the upper index with the newly introduced index of differentiation. The divergence of an arbitrary tensor is defined in precisely the same way:

divergence of a tensor field

Definition 36.2.22 Given a tensor field $\mathbf{T}$, define its **divergence** $\nabla \cdot \mathbf{T}$ to be the tensor obtained from $\nabla \mathbf{T}$ by contracting the last upper index with the covariant derivative index. In components,

$$(\nabla \cdot \mathbf{T})^{i_1 \dots i_{r-1}}_{j_1 \dots j_s} = T^{i_1 \dots i_{r-1} k}_{j_1 \dots j_s ; k}.$$

Example 36.2.23 There is a useful relation between the covariant and the Lie derivative that we derive now. First, let $\mathbf{T}$ be of type $(2, 0)$ and write it in some frame as $\mathbf{T} = T^{ij} \mathbf{e}_i \otimes \mathbf{e}_j$. Apply the covariant derivative with respect to $\mathbf{u}$ to both sides to obtain

$$\nabla_{\mathbf{u}} \mathbf{T} = \mathbf{u}\big(T^{ij}\big) \mathbf{e}_i \otimes \mathbf{e}_j + T^{ij} (\nabla_{\mathbf{u}} \mathbf{e}_i) \otimes \mathbf{e}_j + T^{ij} \mathbf{e}_i \otimes (\nabla_{\mathbf{u}} \mathbf{e}_j).$$

Similarly,

derivation of the relation between the Lie and the covariant derivatives

$$L_{\mathbf{u}} \mathbf{T} = \mathbf{u}\big(T^{ij}\big) \mathbf{e}_i \otimes \mathbf{e}_j + T^{ij} (L_{\mathbf{u}} \mathbf{e}_i) \otimes \mathbf{e}_j + T^{ij} \mathbf{e}_i \otimes (L_{\mathbf{u}} \mathbf{e}_j).$$

Now use $L_{\mathbf{u}} \mathbf{e}_j = [\mathbf{u}, \mathbf{e}_j] = \nabla_{\mathbf{u}} \mathbf{e}_j - \nabla_{\mathbf{e}_j} \mathbf{u}$ to get

$$L_{\mathbf{u}} \mathbf{T} = \nabla_{\mathbf{u}} \mathbf{T} - T^{ij} \big[(\nabla_{\mathbf{e}_i} \mathbf{u}) \otimes \mathbf{e}_j + \mathbf{e}_i \otimes (\nabla_{\mathbf{e}_j} \mathbf{u})\big]. \tag{36.37}$$

On the other hand, if we apply $\nabla_{\mathbf{u}}$ and $L_{\mathbf{u}}$ to both sides of $\delta^i_j = \langle \boldsymbol{\epsilon}^i, \mathbf{e}_j \rangle$ and use $[\mathbf{u}, \mathbf{e}_i] = \nabla_{\mathbf{u}} \mathbf{e}_i - \nabla_{\mathbf{e}_i} \mathbf{u}$, we obtain

$$\nabla_{\mathbf{u}} \boldsymbol{\epsilon}^i = L_{\mathbf{u}} \boldsymbol{\epsilon}^i - (\nabla_{\mathbf{e}_j} \mathbf{u})^i \boldsymbol{\epsilon}^j.$$

It follows that for $\mathbf{T} = T_{ij} \boldsymbol{\epsilon}^i \otimes \boldsymbol{\epsilon}^j$, we have

$$L_{\mathbf{u}} \mathbf{T} = \nabla_{\mathbf{u}} \mathbf{T} + T_{ij} \big[(\nabla_{\mathbf{e}_k} \mathbf{u})^i \boldsymbol{\epsilon}^k \otimes \boldsymbol{\epsilon}^j + (\nabla_{\mathbf{e}_k} \mathbf{u})^j \boldsymbol{\epsilon}^i \otimes \boldsymbol{\epsilon}^k\big]. \tag{36.38}$$

One can use Eqs. (36.37) and (36.38) to generalize to a tensor of type (r, s).

Example 36.2.24 Let f be a function on M. Then ∇f is a one form. Call it ϕ. In a local coordinate system, it can be written as $\phi = \phi_i dx^i$, where

$$\phi_i = \phi(\partial_i) = \nabla f(\partial_i) = \nabla_{\partial_i} f = \partial_i f \equiv \frac{\partial f}{\partial x^i}.$$

Noting that $(\nabla f)_i \equiv f_{;i}$, $\nabla f = f_{;i} dx^i$, and $(\nabla^2 f)_{ij} = f_{;ij}$, let us first find the covariant derivative of ϕ. $\nabla \phi$ is a 2-form, whose components can be found as follows:

$$\begin{aligned}
(\nabla \phi)_{ij} &= \nabla \phi(\partial_i, \partial_j) = \nabla_{\partial_j} \big(\phi(\partial_i)\big) - \phi(\nabla_{\partial_i} \partial_j) \\
&= \partial_j(\phi_i) - \phi\big(\Gamma^k{}_{ij} \partial_k\big) = \partial_j(\partial_i f) - \Gamma^k{}_{ij} \phi_k \\
&= \frac{\partial^2 f}{\partial x^j \partial x^i} - \Gamma^k{}_{ij} \partial_k f = \frac{\partial^2 f}{\partial x^j \partial x^i} - \Gamma^k{}_{ij} \frac{\partial f}{\partial x^k},
\end{aligned}$$

where we used Eq. (36.6). We rewrite this as

$$f_{;ij} = \frac{\partial^2 f}{\partial x^j \partial x^i} - \Gamma^k{}_{ij} \frac{\partial f}{\partial x^k}. \tag{36.39}$$

Reversing the order of indices, we get

$$f_{;ji} = \frac{\partial^2 f}{\partial x^i \partial x^j} - \Gamma^k{}_{ji} \frac{\partial f}{\partial x^k}. \tag{36.40}$$

Subtracting (36.39) from (36.40) and using (36.23), we obtain

$$f_{;ij} - f_{;ji} = \left(\Gamma^k{}_{ij} - \Gamma^k{}_{ji}\right) \frac{\partial f}{\partial x^k} = T^k_{ij} \partial_k f. \tag{36.41}$$

Thus, only if the torsion tensor vanishes are the mixed "partial" covariant derivatives equal.

Now we want to find the difference between the mixed second "partial" covariant derivatives of a vector field $\mathbf{Z} = \xi^k \partial_k$. It is more instructive to use general vectors and then specialize to coordinate vector fields. We are thus interested in $\nabla^2 \mathbf{Z}(\mathbf{Y}, \mathbf{X}) - \nabla^2 \mathbf{Z}(\mathbf{X}, \mathbf{Y})$. Let $\psi \equiv \nabla \mathbf{Z}$. From Eq. (36.6), we have

$$\begin{aligned} \nabla \psi(\mathbf{X}, \mathbf{Y}) &= \nabla_Y \big(\psi(\mathbf{X})\big) - \psi(\nabla_Y \mathbf{X}) \\ &= \nabla_Y \big(\nabla \mathbf{Z}(\mathbf{X})\big) - \nabla \mathbf{Z}(\nabla_Y \mathbf{X}) \\ &= \nabla_Y \nabla_X \mathbf{Z} - \nabla_{\nabla_Y \mathbf{X}} \mathbf{Z}. \end{aligned}$$

Switching $\mathbf{X}$ and $\mathbf{Y}$ and subtracting, we get

$$\begin{aligned} \nabla^2 \mathbf{Z}(\mathbf{Y}, \mathbf{X}) - \nabla^2 \mathbf{Z}(\mathbf{X}, \mathbf{Y}) &= \nabla_X \nabla_Y \mathbf{Z} - \nabla_Y \nabla_X \mathbf{Z} + \nabla_{\nabla_Y \mathbf{X}} \mathbf{Z} - \nabla_{\nabla_X \mathbf{Y}} \mathbf{Z} \\ &= [\nabla_X, \nabla_Y] \mathbf{Z} - \nabla_{\nabla_X \mathbf{Y} - \nabla_Y \mathbf{X}} \mathbf{Z} \\ &= [\nabla_X, \nabla_Y] \mathbf{Z} - \nabla_{\mathsf{T}(\mathbf{X}, \mathbf{Y}) + [\mathbf{X}, \mathbf{Y}]} \mathbf{Z}, \end{aligned}$$

where we used Theorem 36.2.16. We thus have

$$\nabla^2 \mathbf{Z}(\mathbf{Y}, \mathbf{X}) - \nabla^2 \mathbf{Z}(\mathbf{X}, \mathbf{Y}) = [\nabla_X, \nabla_Y] \mathbf{Z} - \nabla_{[\mathbf{X}, \mathbf{Y}]} \mathbf{Z} - \nabla_{\mathsf{T}(\mathbf{X}, \mathbf{Y})} \mathbf{Z},$$

or, using Theorem 36.2.16 again,

$$\nabla^2 \mathbf{Z}(\mathbf{Y}, \mathbf{X}) - \nabla^2 \mathbf{Z}(\mathbf{X}, \mathbf{Y}) = \mathsf{R}(\mathbf{X}, \mathbf{Y}) \mathbf{Z} + \nabla_{\mathsf{T}(\mathbf{Y}, \mathbf{X})} \mathbf{Z}. \tag{36.42}$$

Substituting ∂_i and ∂_j for $\mathbf{Y}$ and $\mathbf{X}$ in the equation above, we can get

$$\xi^i_{;lk} - \xi^i_{;kl} = R^i_{jkl} \xi^j + T^j_{lk} \xi^i_{;j}. \tag{36.43}$$

We leave this as an exercise for the reader.

36.3 Geodesics

Let γ be a curve in M. Denote $\gamma(t)$ by x_t, so that $\gamma(0) = x_0 \in M$. Let γ_s^t be the parallel displacement along the curve γ in M from $T_{x_t}(M)$ to $T_{x_s}(M)$. In particular, consider γ_0^t, the parallel displacement from $T_{x_t}(M)$ to $T_{x_0}(M)$ along γ. It is natural to associate the zero vector in $T_{x_t}(M)$ with x_t.[3] As t varies, the zero vector also varies, and by the parallel displacement γ_0^t, one can monitor how the image of x_t "develops" in $T_{x_0}(M)$.

Definition 36.3.1 The **development** of the curve γ in M into $T_{x_0}(M)$ is the curve $\gamma_0^t(x_t)$ in $T_{x_0}(M)$.

development of a curve

The following theorem, whose proof can be found in [Koba 63, pp 131–132], relates the tangent to the development of a curve and the parallel displacement of its tangent.

Theorem 36.3.2 *Let γ be a curve in M and $\mathbf{Y}_t = \gamma_0^t(\dot{x}_t)$. Let $C_t = \gamma_0^t(x_t)$ be the development of γ in M into $T_{x_0}(M)$. Then*

$$\frac{dC}{dt} = \mathbf{Y}_t.$$

This theorem states that the tangent to the development of a curve is the same as the parallel displacement of the tangent to the curve. In other words, γ_0^t "develops" not only the curve, but its tangent at every point of the curve.

An interesting consequence of this theorem is that if $\dot{x}_t$ is parallel along γ, then $\mathbf{Y}_t$ is independent of t, i.e., $\mathbf{Y}_t$ is constant, say $\mathbf{Y}_t = \mathbf{a}$. Then, $C_t = \mathbf{a}t + \mathbf{b}$. Hence, we have

Corollary 36.3.3 *The development of γ in M into $T_{x_0}(M)$ is a straight line iff $\dot{x}_t$ is parallel along γ.*

Curves in manifolds with a given linear connection bend for two reasons: one is because the curve itself goes back and forth in the manifold; the other is the inherent bending of the manifold itself. The straightest possible lines in a manifold are those which bend only because of the inherent bending of the manifold. Given any curve, we can gauge its bending by parallel displacement of vector fields along that curve. If the vector field has a vanishing covariant derivative, it is said to be parallel along the curve. However, that says nothing about how "curvy" the curve itself is.

To get further insight, let's look at the familiar flat space. In the flat space of a large sheet of paper, construction of a straight line in a given direction starting at a given point P_0 is done by laying down the end of a vector (a straight edge) at P_0 pointing in the given direction, connecting P_0 to a

[3] This association becomes plausible if one specializes to two dimensions and notes that the plane $T_{x_t}(M)$ touches M at x_t, the natural origin of the plane.

neighboring point P_1 along the vector, moving the vector *parallel to itself* to P_1, connecting P_1 to a neighboring point P_2, and continuing the process. In the language of the machinery of the covariant derivative, we might say that a straight line is constructed by transporting the tangent vector parallel to itself.

geodesics defined

Definition 36.3.4 Let M be a manifold and γ a curve in M. Then γ is called a **geodesic** of M if the tangent vector $\dot{x}_t$ at every point of γ is parallel displaced along the curve: $\nabla_{\dot{x}_t}\dot{x}_t = 0$.

affine parameter

Since the definition is in terms of the parameter t, the parametrization of the curve becomes important. Such a parameter, if it exists is called an **affine parameter**.

geodesic equation

It follows from Eq. (36.36)—with $v^k = u^k = dx^k/dt$—that a geodesic curve satisfies the following DE:

$$\frac{d^2x^k}{dt^2} + \Gamma^k{}_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt} = 0. \tag{36.44}$$

This second-order DE, called the **geodesic equation**, will have a unique solution if $x^i(0)$ and $\dot{x}^i(0)$, i.e., the initial point and the initial direction, are given. Thus,

Theorem 36.3.5 *Through a given point and in a given direction passes only one geodesic curve.*

If $s(t)$ is another parametrization of the geodesic curve, then a simple calculation shows that

$$0 = \frac{d^2x^k}{dt^2} + \Gamma^k{}_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt} = \underbrace{\left(\frac{d^2x^k}{ds^2} + \Gamma^k{}_{ij}\frac{dx^i}{ds}\frac{dx^j}{ds}\right)}_{=0}\bigl(s'(t)\bigr)^2 + \frac{dx^k}{ds}s''(t).$$

This requires $s''(t)$ to be zero, or $s = \alpha t + \beta$, with $\alpha, \beta \in \mathbb{R}$.

Corollary 36.3.3 leads immediately to the following:

Proposition 36.3.6 *A curve through $x \in M$ is a geodesic iff its development into $T_x(M)$ is a straight line.*

36.3.1 Riemann Normal Coordinates

Starting with a point P of an n-dimensional manifold M on which a covariant derivative is defined, we can construct a unique geodesic in every direction, i.e., for every vector in $\mathcal{T}_P(M)$. By parallel transportation of the tangent vectors at P, we can construct a vector field in a neighborhood of P: The value of the vector field at Q—assumed to be close enough to P—is the tangent at Q on the geodesic starting at P and passing through Q.[4] The vector field so obtained makes it possible to define an exponential map from the tangent space to the manifold. In fact, the integral curve $\exp(t\mathbf{X})$ of any tangent vector $\mathbf{X}$ in $\mathcal{T}_P(M)$ is simply the geodesic associated with the vector.

The uniqueness of the geodesics establishes a bijection (in fact, a diffeomorphism) between a neighborhood of the origin of $\mathcal{T}_P(M)$ and a neighborhood of P in M. This diffeomorphism can be used to assign coordinates to all points in the vicinity of P. Recall that a coordinate is a smooth bijection from M to $\mathbb{R}^n$. Now choose a basis for $\mathcal{T}_P(M)$ and associate the components of $t\mathbf{X}$ in this basis to the points on the geodesic $\exp(t\mathbf{X})$. Specifically, if $\{a^i\}_{i=1}^n$ are the components of $\mathbf{X}$ in the chosen basis, then

Riemann normal coordinates

$$x^i(t) = a^i t, \quad i = 1, 2, \ldots, n,$$

are the so-called **Riemann normal coordinates** (RNCs) of points on the geodesic of $\mathbf{X}$. The geodesic equations in these coordinates become

$$\Gamma^k{}_{ij} a^i a^j = 0 \quad \Rightarrow \quad \Gamma^k{}_{ji} + \Gamma^k{}_{ij} = 0.$$

In particular, if the torsion vanishes, then $\Gamma^k{}_{ji}$ is symmetric in i and j. Hence, we have the following:

Proposition 36.3.7 *The connection coefficients at a point $P \in M$ vanish in the Riemann normal coordinates at P if the torsion vanishes.*

Using Eq. (36.33), we immediately obtain the following:

Corollary 36.3.8 *Let $\mathbf{T}$ be a tensor field on M with components $T^{i_1 \ldots i_r}_{j_1 \ldots j_s}$ with respect to a Riemann normal coordinate system $\{x^i\}$ at P. Then*

$$T^{i_1 \ldots i_r}_{j_1 \ldots j_s; k} = \frac{\partial}{\partial x^k} T^{i_1 \ldots i_r}_{j_1 \ldots j_s}$$

if the torsion vanishes.

Riemann normal coordinates are very useful in establishing tensor equations. This is because two tensors are identical if and only if their components are the same in any coordinate frame. Therefore, to show that two tensors fields are equal, we pick an arbitrary point in M, erect a set of RNCs, and show that the components of the tensors are equal. Since the connection coefficients vanish in an RNC system, and covariant derivatives are the same as ordinary derivatives, tensor manipulations can be simplified considerably.

[4] We are assuming that through any two neighboring points one can always draw a geodesic. For a proof see [Koba 63, pp. 172–175].

For example, the components of the curvature tensor in RNCs are

$$R^i{}_{jkl} = \frac{\partial \Gamma^i{}_{lj}}{\partial x^k} - \frac{\partial \Gamma^i{}_{kj}}{\partial x^l}. \tag{36.45}$$

This is *not* a tensor relation—the RHS is not a tensor in a general coordinate system. However, if we establish a relation involving the components of the curvature tensor alone, then that relation will hold in all coordinates, i.e., it is a tensor relation. For instance, from the equation above one immediately obtains

$$R^i{}_{jkl} + R^i{}_{ljk} + R^i{}_{klj} = 0.$$

Since this involves only a tensor, it must hold in all coordinate frames. This is the coordinate expression of Bianchi's first identity of Eq. (36.14).[5]

Example 36.3.9 Differentiate the second equation in (36.23) with respect to x^m and evaluate the result in RNC to get

$$R^i{}_{jkl;m} = R^i{}_{jkl,m} = \Gamma^i{}_{lj,km} - \Gamma^i{}_{kj,lm}.$$

From this relation and $\Gamma^i{}_{jl,km} = \Gamma^i{}_{jl,mk}$, we obtain the coordinate expression of Bianchi's second identity of Theorem 36.2.17:

$$R^i{}_{jkl;m} + R^i{}_{jmk;l} + R^i{}_{jlm;k} = 0 \quad \text{and} \quad R^i{}_{j[kl;m]} = 0. \tag{36.46}$$

In Einstein's general relativity, this identity is the analogue of Maxwell's pair of homogeneous equations: $F_{\alpha\beta,\gamma} + F_{\gamma\alpha,\beta} + F_{\beta\gamma,\alpha} = 0$.

Using Proposition 36.3.7, we establish the following tensor identity, which although derived in Riemann normal coordinates, is true in general.

Corollary 36.3.10 *Let $\boldsymbol{\omega}$ be a differential form on M. If the torsion vanishes, then*

$$d\boldsymbol{\omega} = \mathbb{A}(\nabla\boldsymbol{\omega}),$$

where $\mathbb{A}$ is the antisymmetrizer introduced in Eq. (26.14).

36.4 Problems

36.1 Use the fact that $R_{g*}\mathbf{X}$ is a vector at $pg \in L(M)$ to show that the canonical 1-form of $L(M)$ is a tensorial 1-form of type $(GL(n,\mathbb{R}),\mathbb{R}^n)$.

36.2 Derive the two structure equations of (36.5). Hint: For the second equation, use (34.13) with structure constants coming from Example 29.2.7.

[5] See also Problem 36.6 for both of Bianchi's identities.

36.3 Let $\mathbf{Y}$ be a constant vector in (d) of Proposition 36.2.11 to show that $\nabla_X f = \mathbf{X} f$.

36.4 Derive Eq. (36.6).

36.5 In this problem, you are going to prove Bianchi's second identity in terms of curvature tensor.

(a) Show that

$$D^\omega \boldsymbol{\Omega}(\mathbf{X}^*, \mathbf{Y}^*, \mathbf{Z}^*) = \mathrm{Cyc}\big(\mathbf{X}^*\big(\boldsymbol{\Omega}(\mathbf{Y}^*, \mathbf{Z}^*)\big) - \boldsymbol{\Omega}\big([\mathbf{X}^*, \mathbf{Y}^*], \mathbf{Z}^*\big)\big).$$

(b) Using arguments similar to the text, show that

$$p\big(\mathbf{X}^*\big(\boldsymbol{\Omega}(\mathbf{Y}^*, \mathbf{Z}^*)\big)\big) = \nabla_X \mathbf{R}(\mathbf{Y}, \mathbf{Z})$$

(c) Convince yourself that

$$p\boldsymbol{\Omega}\big([\mathbf{X}^*, \mathbf{Y}^*], \mathbf{Z}^*\big) = \mathbf{R}\big(\pi_*[\mathbf{X}^*, \mathbf{Y}^*], \mathbf{Z}\big).$$

(d) Use $\pi_* = p \circ \boldsymbol{\theta}$ and the fact that

$$\boldsymbol{\theta}[\mathbf{X}^*, \mathbf{Y}^*] = d\boldsymbol{\theta}(\mathbf{X}^*, \mathbf{Y}^*) = \boldsymbol{\Theta}(\mathbf{X}^*, \mathbf{Y}^*)$$

to show that $\pi_*[\mathbf{X}^*, \mathbf{Y}^*] = \mathbf{T}(\mathbf{X}, \mathbf{Y})$.

(e) Put everything together and show that $\mathrm{Cyc}[\nabla_X \mathbf{R}(\mathbf{Y}, \mathbf{Z})] = 0$ when torsion tensor vanishes.

36.6 Derive the coordinate expression for Bianchi's first and second identities of Theorem 36.2.18.

36.7 Use $Y_k^i X_j^k = \delta_j^i$ to show that

$$\partial_k^m Y_j^i \equiv \frac{\partial}{\partial X_m^k} Y_j^i = -Y_k^i Y_j^m$$

36.8 From Eq. (36.18) show that $b^i_{\ jk} = \Gamma^l_{km} Y_l^i X_j^m$. Now use this result to rewrite Eq. (36.16) as (36.19).

36.9 Derive Eq. (36.21).

36.10 Prove the formulas in Eq. (36.23).

36.11 Substituting (36.20) in (36.24) and noting the antisymmetry of the wedge product, derive Eq. (36.25).

36.12 Let $\mathbf{Z} = \xi^k \partial_k$. Show that

$$(\nabla \mathbf{Z})^i_j \equiv \xi^i_{;j} = \frac{\partial \xi^i}{\partial x^j} + \Gamma^i_{\ jk} \xi^k$$

and

$$\xi^i_{;lk} - \xi^i_{;kl} = R^i_{jkl} \xi^j + T^j_{lk} \xi^i_{;j}$$