

# The Tangent Structure

In this chapter we introduce the notions of tangent space and cotangent space of a smooth manifold. The union of the tangent spaces of a given manifold will be given a smooth structure making this union a manifold in its own right, called the tangent bundle. Similarly we introduce the cotangent bundle of a smooth manifold. We then discuss vector fields and their integral curves together with the associated dynamic notions of Lie derivative and Lie bracket. Finally, we define and discuss the notion of a 1-form (or covector field), which is the notion dual to the notion of a vector field. One can integrate 1-forms along curves. Such an integration is called a line integral. We explore the concept of exact 1-forms and nonexact 1-forms and their relation to the question of path independence of line integrals.

## 2.1. The Tangent Space

If  $c : (-\epsilon, \epsilon) \rightarrow \mathbb{R}^N$  is a smooth curve, then it is common to visualize the “velocity vector”  $\dot{c}(0)$  as being based at the point  $p = c(0)$ . It is often desirable to explicitly form a separate  $N$ -dimensional vector space for each point  $p$ , whose elements are to be thought of as being based at  $p$ . One way to do this is to use  $\{p\} \times \mathbb{R}^N$  so that a tangent vector based at  $p$  is taken to be a pair  $(p, v)$  where  $v \in \mathbb{R}^N$ . The set  $\{p\} \times \mathbb{R}^N$  inherits a vector space structure from  $\mathbb{R}^N$  in the obvious way. In this context, we *provisionally* denote  $\{p\} \times \mathbb{R}^N$  by  $T_p \mathbb{R}^N$  and refer to it as the tangent space at  $p$ . If we write  $c(t) = (x^1(t), \dots, x^N(t))$ , then the velocity vector of a curve  $c$  at time<sup>1</sup>  $t = 0$  is  $(p, \frac{dx^1}{dt}(0), \dots, \frac{dx^N}{dt}(0))$ , which is based at  $p = c(0)$ . Ambiguously, both

---

<sup>1</sup>It is common to refer to the parameter  $t$  for a curve as “time”, although it may have nothing to do with physical time in a given situation.

$(p, \frac{dx^1}{dt}(0), \dots, \frac{dx^N}{dt}(0))$  and  $(\frac{dx^1}{dt}(0), \dots, \frac{dx^N}{dt}(0))$  are often denoted by  $\dot{c}(0)$  or  $c'(0)$ . A bit more generally, if  $V$  is a finite-dimensional vector space, then  $V$  is a smooth manifold and the tangent space at  $p \in V$  can be provisionally taken to be the set  $\{p\} \times V$ . We use the notation  $v_p := (p, v)$ . If  $v_p := (p, v)$  is a tangent vector at  $p$ , then  $v$  is called the **principal part** of  $v_p$ .

We have a natural isomorphism between  $\mathbb{R}^N$  and  $T_p\mathbb{R}^N$  given by  $v \mapsto (p, v)$ , for any  $p$ . Of course we also have a natural isomorphism  $T_p\mathbb{R}^N \cong T_q\mathbb{R}^N$  for any pair of points given by  $(p, v) \mapsto (q, v)$ . This is sometimes referred to as **distant parallelism**. Here we see the reason that in the context of calculus on  $\mathbb{R}^N$ , the explicit construction of vectors based at a point is often deemed unnecessary. However, from the point of view of manifold theory, the tangent space at a point is a fundamental construction. We will define the notion of a tangent space at a point of a differentiable manifold, and it will be seen that there is, in general, no canonical way to identify tangent spaces at different points.

Actually, we shall give several (ultimately equivalent) definitions of the tangent space. Let us start with the special case of a submanifold of  $\mathbb{R}^N$ . A tangent vector at  $p$  can be variously thought of as the velocity of a curve, as a direction for a directional derivative, and also as a geometric object which has components that depend in a special way on the coordinates used. Let us explore these aspects in the case of a submanifold of  $\mathbb{R}^N$ . If  $M$  is an  $n$ -dimensional regular submanifold of  $\mathbb{R}^N$ , then a smooth curve  $c : (-\epsilon, \epsilon) \rightarrow M$  is also a smooth curve into  $\mathbb{R}^N$  and  $\dot{c}(0)$  is normally thought of as a vector based at the point  $p = c(0)$ . This vector is tangent to  $M$ . The set of all vectors obtained in this way from curves into  $M$  is an  $n$ -dimensional *subspace* of the tangent space of  $\mathbb{R}^N$  at  $p$  (described above). In this special case, this subspace could play the role of the tangent space of  $M$  at  $p$ . Let us tentatively accept this definition of the tangent space at  $p$  and denote it by  $T_pM$ .

Let  $v_p := (p, v) \in T_pM$ . There are three things we should notice about  $v_p$ . First, there are many different curves  $c : (-\epsilon, \epsilon) \rightarrow M$  with  $c(0) = p$  which all give the same tangent vector  $v_p$ , and there is an obvious equivalence relation among these curves: two curves passing through  $p$  at  $t = 0$  are equivalent if they have the same velocity vector. Already one can see that perhaps this could be turned around so that we can think of a tangent vector as an equivalence class of curves. Curves would be equivalent if they agree infinitesimally in some appropriate sense.

The second thing that we wish to bring out is that a tangent vector can be used to construct a directional derivative operator. If  $v_p = (p, v)$  is a tangent vector in  $T_p\mathbb{R}^N$ , then we have a directional derivative operator at  $p$  which is a map  $C^\infty(\mathbb{R}^N) \rightarrow \mathbb{R}$  given by  $f \mapsto Df(p)v$ . Now if  $v_p$  is tangent

to  $M$ , we would like a similar map  $C^\infty(M) \rightarrow \mathbb{R}$ . If  $f$  is only defined on  $M$ , then we do not have  $Df$  to work with but we can just take our directional derivative to be the map given by

$$D_{v_p} : f \mapsto (f \circ c)'(0),$$

where  $c : I \rightarrow M$  is any curve whose velocity at  $t = 0$  is  $v_p$ . Later we use the abstract properties of such a directional derivative to actually *define* the notion of a tangent vector.

Finally, notice how  $v_p$  relates to charts for the submanifold. If  $(U, \mathbf{y})$  is a chart on  $M$  with  $p \in U$ , then by inverting we obtain a map  $\mathbf{y}^{-1} : V \rightarrow M$ , which we may then think of as a map into the ambient space  $\mathbb{R}^N$ . The map  $\mathbf{y}^{-1}$  parameterizes a portion of  $M$ . For convenience, let us suppose that  $\mathbf{y}^{-1}(0) = p$ . Then we have the “coordinate curves”  $y^i \mapsto \mathbf{y}^{-1}(0, \dots, y^i, \dots, 0)$  for  $i = 1, \dots, n$ . The resulting tangent vectors  $E_i$  at  $p$  have principal parts given by the partial derivatives so that

$$E_i := \left( p, \frac{\partial \mathbf{y}^{-1}}{\partial y^i}(0) \right).$$

It can be shown that  $(E_1, \dots, E_n)$  is a basis for  $T_p M$ . For another coordinate system  $\bar{\mathbf{y}}$  with  $\bar{\mathbf{y}}^{-1}(0) = p$ , we similarly define a basis  $(\bar{E}_1, \dots, \bar{E}_n)$ . If  $v_p = \sum_{i=1}^n a^i E_i = \sum_{i=1}^n \bar{a}^i \bar{E}_i$ , then letting  $a = (a^1, \dots, a^n)$  and  $\bar{a} = (\bar{a}^1, \dots, \bar{a}^n)$ , the chain rule can be used to show that

$$\bar{a} = D(\bar{\mathbf{y}} \circ \mathbf{y}^{-1})|_{\mathbf{y}(p)} a,$$

which is classically written as

$$\bar{a}^i = \sum_{j=1}^n \frac{\partial \bar{y}^i}{\partial y^j} a^j.$$

Both  $(a^1, \dots, a^n)$  and  $(\bar{a}^1, \dots, \bar{a}^n)$  represent the tangent vector  $v_p$ , but with respect to different charts. This is a simple example of a transformation law.

The various definitions for the notion of a tangent vector given below in the general setting will be based in turn on the following three ideas: (1) Equivalence classes of curves through a point. (2) Transformation laws for the components of a tangent vector with respect to various charts. (3) The idea of a “derivation” which is a kind of abstract directional derivative. Of course we will also have to show how to relate these various definitions to see that they are really equivalent.

**2.1.1. Tangent space via curves.** Let  $p$  be a point in a smooth  $n$ -manifold  $M$ . Suppose that we have smooth curves  $c_1$  and  $c_2$  mapping into  $M$ , each with open interval domains containing  $0 \in \mathbb{R}$  and with  $c_1(0) = c_2(0) = p$ . We say that  $c_1$  is tangent to  $c_2$  at  $p$  if for all smooth real-valued functions

$f$  defined on an open neighborhood of  $p$ , we have  $(f \circ c_1)'(0) = (f \circ c_2)'(0)$ . This is an equivalence relation on the set of all such curves. The reader should check that this really is an equivalence relation and also do so when we introduce other simple equivalence relations later. Define a **tangent vector at  $p$**  to be an equivalence class under this relation.

**Notation 2.1.** The equivalence class of  $c$  will be denoted by  $[c]$ , but we also denote tangent vectors by notation such as  $v_p$  or  $X_p$ , etc. Eventually we will often denote tangent vectors simply as  $v, w$ , etc., but for the discussion to follow we reserve these letters without the subscript for elements of  $\mathbb{R}^n$  for some  $n$ .

If  $v_p = [c]$  then we will also write  $\dot{c}(0) = v_p$ . The **tangent space**  $T_pM$  is defined to be the set of all tangent vectors at  $p \in M$ . A simple cut-off function argument shows that  $c_1$  is equivalent to  $c_2$  if and only if  $(f \circ c_1)'(0) = (f \circ c_2)'(0)$  for all globally defined smooth functions  $f : M \rightarrow \mathbb{R}$ .

**Lemma 2.2.**  $c_1$  is tangent to  $c_2$  at  $p$  if and only if  $(f \circ c_1)'(0) = (f \circ c_2)'(0)$  for all  $\mathbb{R}^k$ -valued functions  $f$  defined on an open neighborhood of  $p$ .

**Proof.** If  $f = (f^1, \dots, f^n)$ , then  $(f \circ c_1)'(0) = (f \circ c_2)'(0)$  if and only if  $(f^i \circ c_1)'(0) = (f^i \circ c_2)'(0)$  for  $i = 1, \dots, k$ . Thus  $(f \circ c_1)'(0) = (f \circ c_2)'(0)$  if  $c_1$  is tangent to  $c_2$  at  $p$ . Conversely, let  $g$  be a smooth real-valued function defined on an open neighborhood of  $p$  and consider the map  $f = (g, 0, \dots, 0)$ . Then the equality  $(f \circ c_1)'(0) = (f \circ c_2)'(0)$  implies that  $(g \circ c_1)'(0) = (g \circ c_2)'(0)$ .  $\square$

The definition of tangent space just given is very geometric, but it has one disadvantage. Namely, it is not immediately obvious that  $T_pM$  is a vector space in a natural way. The following principle is used to obtain a vector space structure:

**Proposition 2.3** (Consistent transfer of linear structure). *Suppose that  $S$  is a set and  $\{V_\alpha\}_{\alpha \in A}$  is a family of  $n$ -dimensional vector spaces. Suppose that for each  $\alpha$  we have a bijection  $b_\alpha : V_\alpha \rightarrow S$ . If for every  $\alpha, \beta \in A$  the map  $b_\beta^{-1} \circ b_\alpha : V_\alpha \rightarrow V_\beta$  is a linear isomorphism, then there is a unique vector space structure on the set  $S$  such that each  $b_\alpha$  is a linear isomorphism.*

**Proof.** Define addition in  $S$  by  $s_1 + s_2 := b_\alpha(b_\alpha^{-1}(s_1) + b_\alpha^{-1}(s_2))$ . This definition is independent of the choice of  $\alpha$ . Indeed,

$$\begin{aligned} b_\alpha(b_\alpha^{-1}(s_1) + b_\alpha^{-1}(s_2)) &= b_\alpha[b_\alpha^{-1} \circ b_\beta \circ b_\beta^{-1}(s_1) + b_\alpha^{-1} \circ b_\beta \circ b_\beta^{-1}(s_2)] \\ &= b_\alpha \circ b_\alpha^{-1} \circ b_\beta \left[ b_\beta^{-1}(s_1) + b_\beta^{-1}(s_2) \right] \\ &= b_\beta \left( b_\beta^{-1}(s_1) + b_\beta^{-1}(s_2) \right). \end{aligned}$$

The definition of scalar multiplication is  $a \cdot s := b_\alpha(a b_\alpha^{-1}(s))$ , and this is shown to be independent of  $\alpha$  in a similar way. The axioms of a vector space are satisfied precisely because they are satisfied by each  $V_\alpha$ .  $\square$

We will use the above proposition to show that there is a natural vector space structure on  $T_p M$ . For every chart  $(\mathbf{x}_\alpha, U_\alpha)$  with  $p \in U$ , we have a map  $b_\alpha : \mathbb{R}^n \rightarrow T_p M$  given by  $v \mapsto [\gamma_v]$ , where  $\gamma_v : t \mapsto \mathbf{x}_\alpha^{-1}(\mathbf{x}_\alpha(p) + tv)$  for  $t$  in a sufficiently small but otherwise irrelevant interval containing 0.

**Lemma 2.4.** *For each chart  $(U_\alpha, \mathbf{x}_\alpha)$ , the map  $b_\alpha : \mathbb{R}^n \rightarrow T_p M$  is a bijection and  $b_\beta^{-1} \circ b_\alpha = D(\mathbf{x}_\beta \circ \mathbf{x}_\alpha^{-1})(\mathbf{x}_\alpha(p))$ .*

**Proof.** We have

$$\begin{aligned} (\mathbf{x}_\alpha \circ \gamma_v)'(0) &= \left. \frac{d}{dt} \right|_{t=0} \mathbf{x}_\alpha \circ \mathbf{x}_\alpha^{-1}(\mathbf{x}_\alpha(p) + tv) \\ &= \left. \frac{d}{dt} \right|_{t=0} (\mathbf{x}_\alpha(p) + tv) = v. \end{aligned}$$

Suppose that  $[\gamma_v] = [\gamma_w]$  for  $v, w \in \mathbb{R}^n$ . Then by Lemma 2.2 we have

$$v = (\mathbf{x}_\alpha \circ \gamma_v)'(0) = (\mathbf{x}_\alpha \circ \gamma_w)'(0) = w.$$

This means that  $b_\alpha$  is injective.

Next we show that  $b_\alpha$  is surjective. Let  $[c] \in T_p M$  be represented by  $c : (-\epsilon, \epsilon) \rightarrow M$ . Let  $v := (\mathbf{x}_\alpha \circ c)'(0) \in \mathbb{R}^n$ . Then we have  $b_\alpha(v) = [\gamma_v]$ , where  $\gamma_v : t \mapsto \mathbf{x}_\alpha^{-1}(\mathbf{x}_\alpha(p) + tv)$ . But  $[\gamma_v] = [c]$  since for any smooth  $f$  defined near  $p$  we have

$$\begin{aligned} (f \circ \gamma_v)'(0) &= \left. \frac{d}{dt} \right|_{t=0} f \circ \mathbf{x}_\alpha^{-1}(\mathbf{x}_\alpha(p) + tv) = D(f \circ \mathbf{x}_\alpha^{-1})(\mathbf{x}_\alpha(p)) \cdot v \\ &= D(f \circ \mathbf{x}_\alpha^{-1})(\mathbf{x}_\alpha(p)) \cdot (\mathbf{x}_\alpha \circ c)'(0) = (f \circ c)'(0). \end{aligned}$$

Thus  $b_\alpha$  is surjective. From Lemma 2.2 we see that the map  $[c] \mapsto (\mathbf{x}_\alpha \circ c)'(0)$  is well-defined, and from the above we see that this map is exactly  $b_\alpha^{-1}$ . Thus

$$b_\beta^{-1} \circ b_\alpha(v) = \left. \frac{d}{dt} \right|_{t=0} \mathbf{x}_\beta \circ \mathbf{x}_\alpha^{-1}(\mathbf{x}_\alpha(p) + tv) = D(\mathbf{x}_\beta \circ \mathbf{x}_\alpha^{-1})(\mathbf{x}_\alpha(p))v. \quad \square$$

The above lemma and proposition combine to provide a vector space structure on the set of tangent vectors. Let us temporarily call the tangent space defined above, the **kinematic tangent space** and denote it by  $(T_p M)_{\text{kin}}$ . Thus, if  $\mathcal{C}_p$  is the set of smooth curves  $c$  defined on some open interval containing 0 such that  $c(0) = p$ , then

$$(T_p M)_{\text{kin}} = \mathcal{C}_p / \sim,$$

where the equivalence is as described above.

**Exercise 2.5.** Let  $c_1$  and  $c_2$  be smooth curves mapping into a smooth manifold  $M$ , each with open interval domains containing  $0 \in \mathbb{R}$  and with  $c_1(0) = c_2(0) = p$ . Show that

$$(f \circ c_1)'(0) = (f \circ c_2)'(0)$$

for all smooth  $f$  if and only if the curves  $\mathbf{x} \circ c_1$  and  $\mathbf{x} \circ c_2$  have the same velocity vector in  $\mathbb{R}^n$  for some and hence any chart  $(U, \mathbf{x})$ .

**2.1.2. Tangent space via charts.** Let  $\mathcal{A}$  be the maximal atlas for an  $n$ -manifold  $M$ . For fixed  $p \in M$ , consider the set  $\Gamma_p$  of all triples  $(p, v, (U, \mathbf{x})) \in \{p\} \times \mathbb{R}^n \times \mathcal{A}$  such that  $p \in U$ . Define an equivalence relation on  $\Gamma_p$  by requiring that  $(p, v, (U, \mathbf{x})) \sim (p, w, (V, \mathbf{y}))$  if and only if

$$(2.1) \quad w = D(\mathbf{y} \circ \mathbf{x}^{-1})|_{\mathbf{x}(p)} \cdot v.$$

In other words, the derivative at  $\mathbf{x}(p)$  of the coordinate change  $\mathbf{y} \circ \mathbf{x}^{-1}$  “identifies”  $v$  with  $w$ . The set  $\Gamma_p/\sim$  of equivalence classes can be given a vector space structure as follows: For each chart  $(U, \mathbf{x})$  containing  $p$ , we have a map  $b_{(U, \mathbf{x})} : \mathbb{R}^n \rightarrow \Gamma_p/\sim$  given by  $v \mapsto [p, v, (U, \mathbf{x})]$ , where  $[p, v, (U, \mathbf{x})]$  denotes the equivalence class of  $(p, v, (U, \mathbf{x}))$ . To see that this map is a bijection, notice that if  $[p, v, (U, \mathbf{x})] = [p, w, (U, \mathbf{x})]$ , then

$$v = D(\mathbf{x} \circ \mathbf{x}^{-1})|_{\mathbf{x}(p)} \cdot v = w$$

by definition. By Proposition 2.3 we obtain a vector space structure on  $\Gamma_p/\sim$  whose elements are **tangent vectors**. This is another version of the tangent space at  $p$ , and we shall (temporarily) denote this by  $(T_p M)_{\text{phys}}$ . The subscript “phys” refers to the fact that this version of the tangent space is based on a “transformation law” and corresponds to a way of looking at things that has traditionally been popular among physicists. If  $v_p = [p, v, (U, \mathbf{x})] \in (T_p M)_{\text{phys}}$ , then we say that  $v \in \mathbb{R}^n$  represents  $v_p$  with respect to the chart  $(U, \mathbf{x})$ .

This viewpoint takes on a more familiar appearance if we use a more classical notation. Let  $(U, \mathbf{x})$  and  $(V, \mathbf{y})$  be two charts containing  $p$  in their domains. If an  $n$ -tuple  $(v^1, \dots, v^n)$  represents a tangent vector at  $p$  from the point of view of  $(U, \mathbf{x})$ , and if the  $n$ -tuple  $(w^1, \dots, w^n)$  represents the same vector from the point of view of  $(V, \mathbf{y})$ , then (2.1) is expressed in the form

$$(2.2) \quad w^i = \sum_{j=1}^n \frac{\partial y^i}{\partial x^j} \Big|_{\mathbf{x}(p)} v^j,$$

where we write the change of coordinates as  $y^i = y^i(x^1, \dots, x^n)$  with  $1 \leq i \leq n$ .

**Notation 2.6.** It is sometimes convenient to index the maximal atlas:  $\mathcal{A} = \{(U_\alpha, \mathbf{x}_\alpha)\}_{\alpha \in \mathcal{A}}$ . Then we would consider triples of the form  $(p, v, \alpha)$  and let

the defining equivalence relation for  $(T_p M)_{\text{phys}}$  be  $(p, v, \alpha) \sim (p, w, \beta)$  if and only if

$$D(\mathbf{x}_\beta \circ \mathbf{x}_\alpha^{-1})|_{\mathbf{x}_\alpha(p)} \cdot v = w.$$

**2.1.3. Tangent space via derivations.** We abstract the notion of directional derivative for our next approach to the tangent space. There are actually at least two common versions of this approach, and we explain both. Let  $M$  be a smooth manifold of dimension  $n$ . A **tangent vector**  $v_p$  at  $p$  is a linear map  $v_p : C^\infty(M) \rightarrow \mathbb{R}$  with the property that for  $f, g \in C^\infty(M)$ ,

$$v_p(fg) = g(p)v_p(f) + f(p)v_p(g).$$

This is the **Leibniz law**. We may say that a tangent vector at  $p$  is a **derivation** of the algebra  $C^\infty(M)$  with respect to the evaluation map  $\text{ev}_p$  at  $p$  defined by  $\text{ev}_p(f) := f(p)$ . Alternatively, we say that  $v_p$  is a **derivation at  $p$** . The set of such derivations at  $p$  is easily seen to be a vector space which is called the **tangent space** at  $p$  and is denoted by  $T_p M$ . We temporarily distinguish this version of the tangent space from  $(T_p M)_{\text{kin}}$  and  $(T_p M)_{\text{phys}}$  defined previously by denoting it  $(T_p M)_{\text{alg}}$  and referring to it as the **algebraic tangent space**. We could also consider the vector space of derivations of  $C^r(M)$  at a point for  $r < \infty$ , but this would not give a finite-dimensional vector space and so is not a good candidate for the definition of the tangent space (see Problem 18). Recall that if  $(U, \mathbf{x})$  is a chart on an  $n$ -manifold  $M$ , we have defined  $\partial f / \partial x^i$  by

$$\frac{\partial f}{\partial x^i}(p) := D_i(f \circ \mathbf{x}^{-1})(\mathbf{x}(p))$$

(see Definition 1.57).

**Definition 2.7.** Given  $(U, \mathbf{x})$  and  $p$  as above, define the operator  $\frac{\partial}{\partial x^i} \Big|_p : C^\infty(M) \rightarrow \mathbb{R}$  by

$$\frac{\partial}{\partial x^i} \Big|_p f := \frac{\partial f}{\partial x^i}(p).$$

It is often helpful to use the easily verified fact that if  $c^\sharp : (-\epsilon, \epsilon) \rightarrow M$  is the curve defined for sufficiently small  $\epsilon$  by

$$c_i(t) := \mathbf{x}^{-1}(\mathbf{x}(p) + t\mathbf{e}_i),$$

where  $\mathbf{e}_i$  is the  $i$ -th member of the standard basis of  $\mathbb{R}^n$ , then

$$\frac{\partial}{\partial x^i} \Big|_p f = \lim_{h \rightarrow 0} \frac{f(c_i(h)) - f(p)}{h}.$$

From the usual product rule it follows that  $\frac{\partial}{\partial x^i} \Big|_p$  is a derivation at  $p$  and so is an element of  $(T_p M)_{\text{alg}}$ . We will show that  $(\frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p)$  is a basis for the vector space  $(T_p M)_{\text{alg}}$ .

**Lemma 2.8.** *Let  $v_p \in (T_p M)_{\text{alg}}$ . Then*

- (i) *if  $f, g \in C^\infty(M)$  are equal on some neighborhood of  $p$ , then  $v_p(f) = v_p(g)$ ;*
- (ii) *if  $h \in C^\infty(M)$  is constant on some neighborhood of  $p$ , then  $v_p(h) = 0$ .*

**Proof.** (i) Since  $v_p$  is a linear map, it suffices to show that if  $f = 0$  on a neighborhood  $U$  of  $p$ , then  $v_p(f) = 0$ . Of course  $v_p(0) = 0$ . Let  $\beta$  be a cut-off function with support in  $U$  and  $\beta(p) = 1$ . Then we have that  $\beta f$  is identically zero and so

$$\begin{aligned} 0 &= v_p(\beta f) = f(p)v_p(\beta) + \beta(p)v_p(f) \\ &= v_p(f) \quad (\text{since } \beta(p) = 1 \text{ and } f(p) = 0). \end{aligned}$$

(ii) From what we have just shown, it suffices to assume that  $h$  is equal to a constant  $c$  globally on  $M$ . In the special case  $c = 1$ , we have

$$v_p(1) = v_p(1 \cdot 1) = 1 \cdot v_p(1) + 1 \cdot v_p(1) = 2v_p(1),$$

so that  $v_p(1) = 0$ . Finally we have  $v_p(c) = v_p(1c) = c(v_p(1)) = 0$ .  $\square$

**Notation 2.9.** We shall often write  $v_p f$  or  $v_p \cdot f$  in place of  $v_p(f)$ .

We must now deal with a technical issue. We anticipate that the action of a derivation is really a differentiation and so it seems that a derivation at  $p$  should be able to act on a function defined only in some neighborhood  $U$  of  $p$ . It is pretty easy to see how this would work for  $\frac{\partial}{\partial x^i} \Big|_p$ . But the domain of a derivation as defined is the ring  $C^\infty(M)$  and not  $C^\infty(U)$ . There is nothing in the definition that immediately allows an element of  $(T_p M)_{\text{alg}}$  to act on  $C^\infty(U)$  unless  $U = M$ . It turns out that we can in fact identify  $(T_p U)_{\text{alg}}$  with  $(T_p M)_{\text{alg}}$ , and the following discussion shows how this is done. Once we reach a fuller understanding of the tangent space, this identification will be natural and automatic. So, let  $p \in U \subset M$  with  $U$  open. We construct a rather obvious map  $\Phi : (T_p U)_{\text{alg}} \rightarrow (T_p M)_{\text{alg}}$  by using the restriction map  $C^\infty(M) \rightarrow C^\infty(U)$ . For each  $w_p \in T_p U$ , we define  $\widetilde{w}_p : C^\infty(M) \rightarrow \mathbb{R}$  by  $\widetilde{w}_p(f) := w_p(f|_U)$ . It is easy to show that  $\widetilde{w}_p$  is a derivation of the appropriate type and so  $\widetilde{w}_p \in (T_p M)_{\text{alg}}$ . Thus we get a linear map  $\Phi : (T_p U)_{\text{alg}} \rightarrow (T_p M)_{\text{alg}}$ . We want to show that this map is an isomorphism, but notice that we have not yet established the finite-dimensionality of either  $(T_p U)_{\text{alg}}$  or  $(T_p M)_{\text{alg}}$ . First we show that  $\Phi : w_p \mapsto \widetilde{w}_p$  has trivial kernel. So suppose that  $\widetilde{w}_p = 0$ , i.e.  $\widetilde{w}_p(f) = 0$  for all  $f \in C^\infty(M)$ . Let  $h \in C^\infty(U)$ . Pick a cut-off function  $\beta$  with support in  $U$  so that  $\beta h$  extends by zero to a smooth function  $f$  on all of  $M$  that agrees with  $h$  on a neighborhood of



$p$ . Then by the above lemma,  $w_p(h) = w_p(f|_U) = \widetilde{w}_p(f) = 0$ . Thus, since  $h$  was arbitrary, we see that  $w_p = 0$  and so  $\Phi$  has trivial kernel.

Next we show that  $\Phi$  is *onto*. Let  $v_p \in (T_pM)_{\text{alg}}$ . We wish to define  $w_p \in (T_pU)_{\text{alg}}$  by  $w_p(h) := v_p(\beta h)$ , where  $\beta$  is as above and  $\beta h$  is extended by zero to a function in  $C^\infty(M)$ . If  $\beta_1$  is another similar choice of cut-off function, then  $\beta h$  and  $\beta_1 h$  (both extended to all of  $M$ ) agree on a neighborhood of  $p$ , and so by Lemma 2.8,  $v_p(\beta h) = v_p(\beta_1 h)$ . Thus  $w_p$  is well-defined. Thinking of  $\beta(f|_U)$  as defined on  $M$ , we have  $\widetilde{w}_p(f) := w_p(f|_U) = v_p(\beta f|_U) = v_p(f)$  since  $\beta f|_U$  and  $f$  agree on a neighborhood of  $p$ . Thus  $\Phi : (T_pU)_{\text{alg}} \rightarrow (T_pM)_{\text{alg}}$  is an isomorphism.

Because of this isomorphism, we tend to identify  $(T_pU)_{\text{alg}}$  with  $(T_pM)_{\text{alg}}$  and in particular, if  $(U, \mathbf{x})$  is a chart, we think of the derivations  $\frac{\partial}{\partial x^i}|_p$ ,  $1 \leq i \leq n$  as being simultaneously elements of both  $(T_pU)_{\text{alg}}$  and  $(T_pM)_{\text{alg}}$ . In either case the formula is the same:  $\frac{\partial}{\partial x^i}|_p f = \frac{\partial(f \circ \mathbf{x}^{-1})}{\partial u^i}(\mathbf{x}(p))$ .

Notice that agreeing on a neighborhood of a point is an important relation here and this provides motivation for employing the notion of a germ of a function (Definition 1.68). First we establish the basis theorem:

**Theorem 2.10.** *Let  $M$  be an  $n$ -manifold and  $(U, \mathbf{x})$  a chart with  $p \in U$ . Then the  $n$ -tuple of vectors (derivations)  $\left(\frac{\partial}{\partial x^1}|_p, \dots, \frac{\partial}{\partial x^n}|_p\right)$  is a basis for  $(T_pM)_{\text{alg}}$ . Furthermore, for each  $v_p \in (T_pM)_{\text{alg}}$  we have*

$$v_p = \sum_{i=1}^n v_p(x^i) \frac{\partial}{\partial x^i} \Big|_p .$$

**Proof.** From our discussion above we may assume that  $\mathbf{x}(U)$  is a convex set such as a ball of radius  $\varepsilon$  in  $\mathbb{R}^n$ . By composing with a translation we assume that  $\mathbf{x}(p) = 0$ . This makes no difference for what we wish to prove since  $v_p$  applied to a constant is 0. For any smooth function  $g$  defined on the convex set  $\mathbf{x}(U)$  let

$$g_i(u) := \int_0^1 \frac{\partial g}{\partial u^i}(tu) dt \text{ for all } u \in \mathbf{x}(U).$$

The fundamental theorem of calculus can be used to show that  $g = g(0) + \sum g_i u^i$ . We see that  $g_i(0) = \frac{\partial g}{\partial u^i} \Big|_0$ . For a function  $f \in C^\infty(U)$ , we let  $g := f \circ \mathbf{x}^{-1}$ . Using the above, we arrive at the expression  $f = f(p) + \sum f_i x^i$ , and applying  $\frac{\partial}{\partial x^i} \Big|_p$  we get  $f_i(p) = \frac{\partial f}{\partial x^i} \Big|_p$ . Now apply the derivation  $v_p$  to

$f = f(p) + \sum f_i x^i$  to obtain

$$\begin{aligned} v_p f &= 0 + \sum v_p(f_i x^i) \\ &= \sum v_p(x^i) f_i(p) + \sum 0 v_p f_i \\ &= \sum v_p(x^i) \left. \frac{\partial f}{\partial x^i} \right|_p. \end{aligned}$$

This shows that  $v_p = \sum v_p(x^i) \left. \frac{\partial}{\partial x^i} \right|_p$  and thus we have a spanning set. To see that  $(\left. \frac{\partial}{\partial x^1} \right|_p, \dots, \left. \frac{\partial}{\partial x^n} \right|_p)$  is a linearly independent set, let us assume that  $\sum a^i \left. \frac{\partial}{\partial x^i} \right|_p = 0$  (the zero derivation). Applying this to  $x^j$  gives  $0 = \sum a^i \left. \frac{\partial x^j}{\partial x^i} \right|_p = \sum a^i \delta_i^j = a^j$ , and since  $j$  was arbitrary, we get the result.  $\square$

**Remark 2.11.** On the manifold  $\mathbb{R}^n$ , we have the identity map  $\text{id} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  which gives the standard chart. As is often the case, the simplest situations have the most confusing notation because of the various identifications that may exist. On  $\mathbb{R}$ , there is one coordinate function, which we often denote by either  $u$  or  $t$ . This single function is just  $\text{id}_{\mathbb{R}}$ . The basis vector at  $t_0 \in \mathbb{R}$  associated to this coordinate is  $\left. \frac{\partial}{\partial u} \right|_{t_0}$  (or  $\left. \frac{\partial}{\partial t} \right|_{t_0}$ ). If we think of the tangent space at  $t_0 \in \mathbb{R}$  as being  $\{t_0\} \times \mathbb{R}$ , then  $\left. \frac{\partial}{\partial u} \right|_{t_0}$  is just  $(t_0, 1)$ . It is also common to denote  $\left. \frac{\partial}{\partial u} \right|_{t_0}$  by “1” regardless of the point  $t_0$ .

Above, we used the notion of a derivation as one way to define a tangent vector. There is a slight variation of this approach that allows us to worry a bit less about the relation between  $(T_p U)_{\text{alg}}$  and  $(T_p M)_{\text{alg}}$ . Let  $\mathcal{F}_p = C_p^\infty(M, \mathbb{R})$  be the algebra of germs of functions defined near  $p$ . Recall that if  $f$  is a representative for the equivalence class  $[f] \in \mathcal{F}_p$ , then we can unambiguously define the value of  $[f]$  at  $p$  by  $[f](p) = f(p)$ . Thus we have an evaluation map  $\text{ev}_p : \mathcal{F}_p \rightarrow \mathbb{R}$ .

**Definition 2.12.** A **derivation** (with respect to the evaluation map  $\text{ev}_p$ ) of the algebra  $\mathcal{F}_p$  is a map  $\mathcal{D}_p : \mathcal{F}_p \rightarrow \mathbb{R}$  such that  $\mathcal{D}_p([f][g]) = f(p)\mathcal{D}_p[g] + g(p)\mathcal{D}_p[f]$  for all  $[f], [g] \in \mathcal{F}_p$ .

The set of all these derivations on  $\mathcal{F}_p$  is easily seen to be a real vector space and is sometimes denoted by  $\text{Der}(\mathcal{F}_p)$ .

**Remark 2.13.** The notational distinction between a function and its germ at a point is not always maintained;  $\mathcal{D}_p f$  is taken to mean  $\mathcal{D}_p[f]$ .

Let  $M$  be a smooth manifold of dimension  $n$ . Consider the set of all germs of  $C^\infty$  functions  $\mathcal{F}_p$  at  $p \in M$ . The vector space  $\text{Der}(\mathcal{F}_p)$  of derivations of  $\mathcal{F}_p$  with respect to the evaluation map  $\text{ev}_p$  could also be taken as the definition of the tangent space at  $p$ . This would be a slight variation of what we have called the algebraic tangent space.

## 2.2. Interpretations

We will now show how to move from one definition of tangent vector to the next. Let  $M$  be a (smooth)  $n$ -manifold. Consider a tangent vector  $v_p$  as an equivalence class of curves represented by  $c : I \rightarrow M$  with  $c(0) = p$ . We obtain a derivation by defining

$$v_p f := \left. \frac{d}{dt} \right|_{t=0} f \circ c.$$

This gives a map  $(T_p M)_{\text{kin}} \rightarrow (T_p M)_{\text{alg}}$  which can be shown to be an isomorphism. We also have a natural isomorphism  $(T_p M)_{\text{kin}} \rightarrow (T_p M)_{\text{phys}}$ . Given  $[c] \in (T_p M)_{\text{kin}}$ , we obtain an element  $v_p \in (T_p M)_{\text{phys}}$  by letting  $v_p$  be the equivalence class of the triple  $(p, v, (U, \mathbf{x}))$ , where  $v^i := \left. \frac{d}{dt} \right|_{t=0} x^i \circ c$  for a chart  $(U, \mathbf{x})$  with  $p \in U$ .

If  $v_p$  is a derivation at  $p$  and  $(U, \mathbf{x})$  an admissible chart with domain containing  $p$ , then  $v_p$ , as a tangent vector in the sense of Definition 2.1.2, is represented by the triple  $(p, v, (U, \mathbf{x}))$ , where  $v = (v^1, \dots, v^n)$  is given by

$$v^i = v_p x^i \quad (v_p \text{ is acting as a derivation}).$$

This gives us an isomorphism  $(T_p M)_{\text{alg}} \rightarrow (T_p M)_{\text{phys}}$ .

Next we exhibit the inverse isomorphism  $(T_p M)_{\text{phys}} \rightarrow (T_p M)_{\text{alg}}$ . Suppose that  $[(p, v, (U, \mathbf{x}))] \in (T_p M)_{\text{phys}}$  where  $v \in \mathbb{R}^n$ . We obtain a derivation by defining

$$v_p f = D(f \circ \mathbf{x}^{-1})|_{\mathbf{x}(p)} \cdot v.$$

In other words,

$$v_p f = \sum_{i=1}^n v^i \left. \frac{\partial}{\partial x^i} \right|_p f$$

for  $v = (v^1, \dots, v^n)$ . It is an easy exercise that  $v_p$  defined in this way is independent of the representative triple  $(p, v, (U, \mathbf{x}))$ .

**We now adopt the explicitly flexible attitude of interpreting a tangent vector in any of the ways we have described above depending on the situation. Thus we effectively identify the spaces  $(T_p M)_{\text{kin}}$ ,  $(T_p M)_{\text{phys}}$  and  $(T_p M)_{\text{alg}}$ . Henceforth we use the notation  $T_p M$  for the tangent space of a manifold  $M$  at a point  $p$ .**

**Definition 2.14.** The dual space to a tangent space  $T_p M$  is called the **cotangent space** and is denoted by  $T_p^* M$ . An element of  $T_p^* M$  is referred to as a **covector**.

The basis for  $T_p^* M$  that is dual to the coordinate basis  $(\left. \frac{\partial}{\partial x^1} \right|_p, \dots, \left. \frac{\partial}{\partial x^n} \right|_p)$  described above is denoted  $(dx^1|_p, \dots, dx^n|_p)$ . By definition  $dx^i|_p \left( \left. \frac{\partial}{\partial x^j} \right|_p \right) = \delta_j^i$ . (Note that  $\delta_j^i = 1$  if  $i = j$  and  $\delta_j^i = 0$  if  $i \neq j$ . The symbols  $\delta_{ij}$  and

$\delta^{ij}$  are defined similarly.) The reason for the differential notation  $dx^i$  will be explained below. Sometimes one abbreviates  $\frac{\partial}{\partial x^j}\Big|_p$  and  $dx^i\Big|_p$  to  $\frac{\partial}{\partial x^j}$  and  $dx^i$  respectively, but there is some risk of confusion since later  $\frac{\partial}{\partial x^j}$  and  $dx^i$  will more properly denote not elements of the vector spaces  $T_pM$  and  $T_p^*M$ , but rather *fields* defined over a chart domain. More on this shortly.

**2.2.1. Tangent space of a vector space.** Our provisional definition of the tangent space at point  $p$  in a vector space  $V$  was the set  $\{p\} \times V$ , but this set does not immediately fit any of the definitions of tangent space just given. This is remedied by finding a natural isomorphism  $\{p\} \times V \cong T_pV$ . One may pick a version of the tangent space and then exhibit a natural isomorphism directly, but we take a slightly different approach. Namely, we first define a natural map  $j_p : V \rightarrow T_pV$ . We think in terms of equivalence classes of curves. For each  $v \in V$ , let  $c_{p,v} : \mathbb{R} \rightarrow V$  be the curve  $c_{p,v}(t) := p + tv$ . Then

$$j_p(v) := [c_{p,v}] \in T_pV.$$

As a derivation,  $j_p(v)$  acts according to

$$f \mapsto \left. \frac{d}{dt} \right|_0 f(p + tv).$$

On the other hand, we have the obvious projection  $\text{pr}_2 : \{p\} \times V \rightarrow V$ . Then our natural isomorphism  $\{p\} \times V \cong T_pV$  is just  $j_p \circ \text{pr}_2$ . The isomorphism between the vector spaces  $\{p\} \times V$  and  $T_pV$  is so natural that they are often identified. Of course, since  $T_pV$  itself has various manifestations ( $(T_pV)_{\text{alg}}$ ,  $(T_pV)_{\text{phys}}$ , and  $(T_pV)_{\text{kin}}$ ), we now have a multitude of spaces which are potentially being identified in the case of a vector space.

**Note:** Because of the identification of  $\{p\} \times V$  with  $T_pV$ , we shall often denote by  $\text{pr}_2$  the map  $T_pV \rightarrow V$ . Furthermore, in certain contexts,  $T_pV$  is identified with  $V$  itself. The potential identifications introduced here are often referred to as “canonical” or “natural” and  $j_p : V \rightarrow T_pV$  is often called the **canonical** or **natural** isomorphism. The inverse map  $T_pV \rightarrow V$  is also referred to as the **canonical** or **natural** isomorphism. Context will keep things straight.

### 2.3. The Tangent Map

The first definition given below of the tangent map at  $p \in M$  of a smooth map  $f : M \rightarrow N$  will be considered our main definition, but the others are actually equivalent. Given  $f$  and  $p$  as above, we wish to define a linear map  $T_p f : T_pM \rightarrow T_{f(p)}N$ . Since we have several definitions of tangent space, we expect to see several equivalent definitions of the tangent map. For the first definition we think of  $T_pM$  as  $(T_pM)_{\text{kin}}$ .

# Connections and Covariant Derivatives

The terms “covariant derivative” and “connection” are sometimes treated as synonymous. In fact, a covariant derivative is sometimes called a Koszul connection. From one point of view, the central idea is that of measuring the rate of change of sections of bundles in the direction of a vector or vector field on the base manifold. Here the derivative viewpoint is prominent. From another related point of view, a connection provides an extra structure that gives a principled way of lifting curves from the base to the total space. The lifts are parallel sections along the curve. In this chapter, we will always take the typical fiber of an  $\mathbb{F}$ -vector bundle of rank  $k$  to be  $\mathbb{F}^k$ . We let  $(\mathbf{e}_1, \dots, \mathbf{e}_k)$  be the standard basis of  $\mathbb{F}^k$ . Thus every vector bundle chart  $(U, \phi)$  is associated with a frame field  $(e_1, \dots, e_k)$  where  $e_i(x) := \phi^{-1}(x, \mathbf{e}_i)$ .

## 12.1. Definitions

Let  $\pi : E \rightarrow M$  be a smooth  $\mathbb{F}$ -vector bundle of rank  $k$  over a manifold  $M$ . A covariant derivative can either be defined as a map  $\nabla : \mathfrak{X}(M) \times \Gamma(M, E) \rightarrow \Gamma(M, E)$  with certain properties from which one derives a well-defined map  $\nabla : TM \times \Gamma(M, E) \rightarrow \Gamma(M, E)$  with nice properties or the other way around. We rather arbitrarily start with the first of these definitions.

**Definition 12.1.** A **covariant derivative** or **Koszul connection** on a smooth  $\mathbb{F}$ -vector bundle  $E \rightarrow M$  is a map  $\nabla : \mathfrak{X}(M) \times \Gamma(M, E) \rightarrow \Gamma(M, E)$  (where  $\nabla(X, s)$  is written as  $\nabla_X s$ ) satisfying the following four properties:

- (i)  $\nabla_{fX}(s) = f\nabla_X s$  for all  $f \in C^\infty(M)$ ,  $X \in \mathfrak{X}(M)$  and  $s \in \Gamma(M, E)$ ;

- (ii)  $\nabla_{X_1+X_2}s = \nabla_{X_1}s + \nabla_{X_2}s$  for all  $X_1, X_2 \in \mathfrak{X}(M)$  and  $s \in \Gamma(M, E)$ ;
- (iii)  $\nabla_X(s_1 + s_2) = \nabla_X s_1 + \nabla_X s_2$  for all  $X \in \mathfrak{X}(M)$  and  $s_1, s_2 \in \Gamma(M, E)$ ;
- (iv)  $\nabla_X(fs) = (Xf)s + f\nabla_X s$  for all  $f \in C^\infty(M; \mathbb{F})$ ,  $X \in \mathfrak{X}(M)$  and  $s \in \Gamma(M, E)$ .

For a fixed  $X \in \mathfrak{X}(M)$ , the map  $\nabla_X : \Gamma(M, E) \rightarrow \Gamma(M, E)$  is called the **covariant derivative with respect to  $X$** .

As we will see below, this definition is enough to imply that  $\nabla$  induces maps  $\nabla^U : \mathfrak{X}(U) \times \Gamma(U, E) \rightarrow \Gamma(U, E)$ , one for each open  $U \subset M$ , that are naturally related in a sense we make precise below (this is not necessarily true for infinite-dimensional manifolds). Furthermore, we also prove that for a fixed  $p \in M$ , the value  $(\nabla_X s)(p)$  depends only on the value of  $X$  at  $p$  and only on the values of  $s$  along any smooth curve  $c$  representing  $X_p$ . Thus we get a well-defined map  $\nabla : TM \times \Gamma(M, E) \rightarrow \Gamma(M, E)$  such that  $\nabla_v s = (\nabla_X s)(p)$  for any extension of  $v \in T_p M$  to a vector field  $X$  with  $X_p = v$ . The resulting properties are

- (i')  $\nabla_{av}(s) = a\nabla_v s$  for all  $a \in \mathbb{R}$ ,  $v \in TM$  and  $s \in \Gamma(M, E)$ ;
- (ii') for all  $p \in M$  we have  $\nabla_{v_1+v_2}s = \nabla_{v_1}s + \nabla_{v_2}s$  for all  $v_1, v_2 \in T_p M$ , and  $s \in \Gamma(M, E)$ ;
- (iii')  $\nabla_v(s_1 + s_2) = \nabla_v s_1 + \nabla_v s_2$  for all  $v \in TM$  and  $s_1, s_2 \in \Gamma(M, E)$ ;
- (iv') for all  $p \in M$  we have  $\nabla_v(fs) = (vf)s(p) + f(p)\nabla_v s$  for all  $v \in T_p M$ ,  $s \in \Gamma(M, E)$  and  $f \in C^\infty(M; \mathbb{F})$ ;
- (v')  $p \mapsto \nabla_{X_p}s$  is smooth for all smooth vector fields  $X$ .

A map satisfying these properties is *also* called a **covariant derivative** (or **Koszul connection**). Note that it is easy to obtain a Koszul connection in the first sense since we just let  $(\nabla_X s)(p) := \nabla_{X_p}s$ .

**Definition 12.2.** A covariant derivative on the tangent bundle  $TM$  of an  $n$ -manifold  $M$  is usually referred to as a covariant derivative on  $M$ .

In Chapter 4 we have already met the Levi-Civita connection  $\overline{\nabla}$  on  $\mathbb{R}^n$ , which is, from the current point of view, a Koszul connection on the tangent bundle of  $\mathbb{R}^n$ . The definition of this connection takes advantage of the natural identification of tangent spaces which makes taking the difference quotient possible:

$$\overline{\nabla}_v X = \lim_{t \rightarrow 0} \frac{X(p + tv) - X(p)}{t}.$$

In that same chapter we obtained, by a projection, a covariant derivative on (the tangent bundle of) any hypersurface in  $\mathbb{R}^n$ . A covariant derivative on

a submanifold of arbitrary codimension can be obtained in the same way. Let  $M$  be a submanifold of  $\mathbb{R}^n$  and let  $X \in \mathfrak{X}(M)$  and  $v \in T_p M$ . We have

$$\nabla_v X := \left( \frac{d}{dt} \Big|_0 X \circ c \right)^\top \in T_p M.$$

One may easily verify that  $\nabla$ , so defined, is a covariant derivative (in the second sense above).

Returning to the case of a general vector bundle, let us consider how covariant differentiation behaves with respect to restriction to open subsets of our manifold. Recall the restriction map  $r_V^U : \Gamma(U, E) \rightarrow \Gamma(V, E)$  given by  $r_V^U : \sigma \mapsto \sigma|_V$  and where  $V \subset U$ .

**Definition 12.3.** A **natural covariant derivative**  $\nabla$  on a smooth  $\mathbb{F}$ -vector bundle  $E \rightarrow M$  is an assignment to each open set  $U \subset M$  of a map  $\nabla^U : \mathfrak{X}(U) \times \Gamma(U, E) \rightarrow \Gamma(U, E)$  written as  $\nabla^U : (X, \sigma) \mapsto \nabla_X^U \sigma$  such that the following assertions hold:

- (i) For every open  $U \subset M$ , the map  $\nabla^U$  is a Koszul connection on the restricted bundle  $E|_U \rightarrow U$ .
- (ii) For nested open sets  $V \subset U$ , we have  $r_V^U(\nabla_X^U \sigma) = \nabla_{r_V^U X}^V r_V^U \sigma$  (naturality with respect to restrictions).
- (iii) For  $X \in \mathfrak{X}(U)$  and  $\sigma \in \Gamma(U, E)$  the value  $(\nabla_X^U \sigma)(p)$  only depends on the value of  $X$  at  $p \in U$ .

Here  $\nabla_X^U \sigma$  is called the covariant derivative of  $\sigma$  with respect to  $X$ . We will denote all of the maps  $\nabla^U$  by the single symbol  $\nabla$  when there is no chance of confusion. We have explicitly worked the naturality conditions (ii) and (iii) into the definition of a natural covariant derivative, so this definition is also appropriate for infinite-dimensional manifolds. The definition of Koszul connection did not specifically include these naturality features and was only defined for global sections. We shall now see that, in the case of finite-dimensional manifolds, a Koszul connection gives a natural covariant derivative for free.

**Lemma 12.4.** *Suppose  $\nabla : \mathfrak{X}(M) \times \Gamma(E) \rightarrow \Gamma(E)$  is a Koszul connection for the vector bundle  $E \rightarrow M$ . Then if for some open  $U$  either  $X|_U = 0$  or  $\sigma|_U = 0$ , then*

$$(\nabla_X \sigma)(p) = 0 \text{ for all } p \in U.$$

**Proof.** We prove the case of  $\sigma|_U = 0$  and leave the case of  $X|_U = 0$  to the reader.

Let  $q \in U$ . Then there is some relatively compact open set  $V$  with  $q \in \overline{V} \subset U$  and a smooth function  $f$  that is identically one on  $V$  and zero

outside of  $U$ . Thus  $f\sigma \equiv 0$  on  $M$  and so since  $\nabla$  is linear, we have  $\nabla(f\sigma) \equiv 0$  on  $M$ . Thus since (iv) of Definition 12.1 holds for global fields, we have

$$\begin{aligned}\nabla_X(f\sigma)(q) &= f(p)(\nabla_X\sigma)(q) + (X_q f)\sigma(q) \\ &= (\nabla_X\sigma)(q) = 0.\end{aligned}$$

Since  $q \in U$  was arbitrary, we have the result.  $\square$

We now define a natural covariant derivative derived from a given Koszul connection  $\nabla$ . Given any open set  $U \subset M$ , we define  $\nabla^U : \mathfrak{X}(U) \times \Gamma(E|_U) \rightarrow \Gamma(E|_U)$  by

$$(12.1) \quad (\nabla_X^U \sigma)(p) := (\nabla_{\tilde{X}} \tilde{\sigma})(p), \quad p \in U,$$

for  $\tilde{X} \in \mathfrak{X}(M)$  and  $\tilde{\sigma} \in \Gamma(E)$  chosen to be any sections which agree with  $X$  and  $\sigma$  on some open  $V$  with  $p \in V \subset \bar{V} \subset U$ . By the above lemma this definition does not depend on the choices of  $\tilde{X}$  and  $\tilde{\sigma}$ .

**Proposition 12.5.** *Let  $E \rightarrow M$  be a rank  $k$  vector bundle and suppose that  $\nabla : \mathfrak{X}(M) \times \Gamma(E) \rightarrow \Gamma(E)$  is a Koszul connection. If for each open  $U$  we define  $\nabla^U$  as in (12.1) above, then the assignment  $U \mapsto \nabla^U$  is a natural covariant derivative as in Definition 12.3.*

**Proof.** We must show that (i), (ii) and (iii) of Definition 12.3 hold. It is easily checked that (i) holds, that is, that  $\nabla^U$  is a Koszul connection for each  $U$ . The demonstration that (ii) holds is also easy and we leave it for the reader to check. Now since  $X \rightarrow \nabla_X \sigma$  is linear over  $C^\infty(U)$ , (iii) follows by familiar arguments ( $\nabla_X \sigma$  is linear over functions in the argument  $X$ ).  $\square$

Because of this last lemma, we may define  $\nabla_{v_p} \sigma$  for  $v_p \in T_p M$  by

$$\nabla_{v_p} \sigma := (\nabla_X \sigma)(p),$$

where  $X$  is any vector field with  $X(p) = v_p \in T_p M$ . We say that  $\nabla_X \sigma$  is “tensorial” in the variable  $X$ . The result can be seen as a special case of Proposition 6.55. We now see that it is safe to write expressions not directly justified by the definition of Koszul connection. For example, if  $X \in \mathfrak{X}(U)$  and  $\sigma \in \Gamma(V, E)$ , where  $U \cap V \neq \emptyset$ , then  $\nabla_X \sigma$  is taken to be an element of  $\Gamma(U \cap V, E)$  defined by

$$(\nabla_X \sigma)(p) = \nabla_{X_p} \sigma := \nabla_{X_p}^U \sigma \text{ for all } p \in U \cap V.$$

This is a particularly useful convention when  $U$  is the domain of a chart  $(U, \mathbf{x})$  and  $X = \frac{\partial}{\partial x^i}$  and also when  $\sigma$  is a member of a frame field of the vector bundle defined on some open set.

In the same way that one extends a derivation on vector fields to a tensor derivation, one may show that a covariant derivative on a vector bundle induces naturally related covariant derivatives on all the multilinear bundles.



In particular, if  $\pi^* : E^* \rightarrow M$  denotes the dual bundle to  $\pi : E \rightarrow M$  we may define connections on  $\pi^* : E^* \rightarrow M$  and on  $\pi \otimes \pi^* : E \otimes E^* \rightarrow M$ . We do this in such a way that for  $s \in \Gamma(M, E)$  and  $s^* \in \Gamma(M, E^*)$  we have

$$\nabla_X^{E \otimes E^*} (s \otimes s^*) = \nabla_X s \otimes s^* + s \otimes \nabla_X^{E^*} s^*,$$

and

$$(\nabla_X^{E^*} s^*)(s) = X(s^*(s)) - s^*(\nabla_X s).$$

Of course, the second formula follows from the requirement that covariant differentiation commutes with contraction:

$$\begin{aligned} X(s^*(s)) &= (\nabla_X C(s \otimes s^*)) = C(\nabla_X^{E \otimes E^*} (s \otimes s^*)) \\ &= C\left(\nabla_X s \otimes s^* + s \otimes \nabla_X^{E^*} s^*\right) = s^*(\nabla_X s) + (\nabla_X^{E^*} s^*)(s), \end{aligned}$$

where  $C$  denotes the contraction given by  $s \otimes \alpha \mapsto \alpha(s)$ . All this works like the tensor derivation extension procedure discussed previously, and we often write all of these covariant derivatives with the single symbol  $\nabla$ .

The bundle  $E \otimes E^* \rightarrow M$  is naturally isomorphic to  $\text{End}(E)$ , and by this isomorphism we get a connection on  $\text{End}(E)$ . If we identify elements of  $\Gamma(\text{End}(E))$  with  $\text{End}(\Gamma E)$  (see Proposition 6.55), then we may use the following formula for the definition of the connection on  $\text{End}(E)$ :

$$(\nabla_X A)(s) := \nabla_X(A(s)) - A(\nabla_X s).$$

Indeed, since  $C : s \otimes A \mapsto A(s)$  is a contraction, we must have

$$\begin{aligned} \nabla_X(A(s)) &= C(\nabla_X s \otimes A + s \otimes \nabla_X A) \\ &= A(\nabla_X s) + (\nabla_X A)(s). \end{aligned}$$

Notice that if we fix  $s \in \Gamma(E)$ , then for each  $p \in M$ , we have an element  $(\nabla s)(p)$  of  $L(T_p M, E_p) \cong E \otimes T^* M$  given by

$$(\nabla s)(p) : v_p \mapsto \nabla_{v_p} s \text{ for all } v_p \in T_p M.$$

Thus we obtain a section  $\nabla s$  of  $E \otimes T^* M$  given by  $p \mapsto (\nabla s)(p)$ , which is easily shown to be smooth. In this way, we can also think of a connection as giving a map

$$\nabla : \Gamma(E) \rightarrow \Gamma(E \otimes T^* M)$$

with the property that

$$\nabla f s = f \nabla s + s \otimes df$$

for all  $s \in \Gamma(E)$  and  $f \in C^\infty(M)$ . Since, by definition,  $\Gamma(E) = \Omega^0(E)$  and  $\Gamma(E \otimes T^* M) = \Omega^1(E)$ , we really have a map  $\Omega^0(E) \rightarrow \Omega^1(E)$ . Later we will extend to a map  $\Omega^k(E) \rightarrow \Omega^{k+1}(E)$  for all integral  $k \geq 0$ . Now if  $X$  is a smooth vector field, then  $X \mapsto \nabla_X s \in \Gamma(E)$ , so we may also view  $\nabla s$  as an element of  $\text{Hom}(\mathfrak{X}(M), \Gamma(E))$ .

## 12.2. Connection Forms

Let  $\pi : E \rightarrow M$  be a rank  $r$  vector bundle with a connection  $\nabla$ . Recall that a choice of a local frame field over an open set  $U \subset M$  is equivalent to a trivialization of the restriction  $\pi_U : E|_U \rightarrow U$ . Namely, if  $\phi = (\pi, \Phi)$  is such a trivialization over  $U$ , then defining  $e_i(x) = \phi^{-1}(x, \mathbf{e}_i)$ , where  $(\mathbf{e}_i)$  is the standard basis of  $\mathbb{F}^n$ , we have a frame field  $(e_1, \dots, e_k)$ . We now examine the expression for a given covariant derivative from the viewpoint of such a local frame. It is not hard to see that for every such frame field there must be a matrix of 1-forms  $\omega = (\omega_j^i)_{1 \leq i, j \leq r}$  such that for  $X \in \Gamma(U, E)$  we may write

$$\nabla_X e_j = \sum_{i=1}^k \omega_j^i(X) e_i.$$

The forms  $\omega_j^i$  are called **connection forms**.

**Proposition 12.6.** *If  $s = \sum_i s^i e_i$  is the local expression of a section  $s \in \Gamma(E)$  in terms of a local frame field  $(e_1, \dots, e_k)$ , then the following local expression holds:*

$$\nabla_X s = \sum_i (X s^i + \sum_r \omega_r^i(X) s^r) e_i.$$

**Proof.** We simply compute:

$$\begin{aligned} \nabla_X s &= \nabla_X \left( \sum_i s^i e_i \right) \\ &= \sum_i (X s^i) e_i + \sum_i s^i \nabla_X e_i \\ &= \sum_i (X s^i) e_i + \sum_{i,j} s^i \omega_i^j(X) e_j \\ &= \sum_i (X s^i) e_i + \sum_{i,r} s^r \omega_r^i(X) e_i \\ &= \sum_i \left( X s^i + \sum_r \omega_r^i(X) s^r \right) e_i. \quad \square \end{aligned}$$

So the  $i$ -th component of  $\nabla_X s$  is

$$(12.2) \quad (\nabla_X s)^i = X s^i + \sum_r \omega_r^i(X) s^r.$$

We may surely choose  $U$  small enough that it is also the domain of a coordinate frame  $\{\partial_\mu\}$  for  $M$ . Thus we have

$$(12.3) \quad \nabla_{\partial_\mu} e_j = \sum_k \omega_{\mu j}^k e_k,$$

where  $\omega_{\mu j}^k = \omega_i^j(\partial_\mu)$ . We now have the local formula

$$(12.4) \quad \nabla_X s = \sum_{i=1}^k \left( \sum_{\mu=1}^n X^\mu \partial_\mu s^i + \sum_{\mu=1}^n \sum_{r=1}^k X^\mu \omega_{\mu r}^i s^r \right) e_i.$$

Or, using the summation convention,

$$\nabla_X s = (X^\mu \partial_\mu s^i + X^\mu \omega_{\mu r}^i s^r) e_i.$$

Now suppose that we have two moving frames whose domains overlap, say  $u = (e_1, \dots, e_k)$  and  $u' = (e'_1, \dots, e'_k)$ . Let us examine how the matrix of 1-forms  $\omega = (\omega_i^j)$  is related to the corresponding  $\omega' = (\omega'^j_i)$  defined in terms of the frame  $u'$ . The change of frame is

$$e'_i = \sum_j g_i^j e_j,$$

which in matrix notation is

$$u' = ug$$

for some smooth  $g : U \cap U' \rightarrow \text{GL}(n)$ . (We treat  $u$  as a row vector of fields.) For a given moving frame, let  $\nabla u := [\nabla e_1, \dots, \nabla e_k]$ . Differentiating both sides of  $u' = ug$  and using matrix notation we have

$$\begin{aligned} u' &= ug, \\ \nabla u' &= \nabla(ug), \\ u' \omega' &= (\nabla u)g + u dg, \\ u' \omega' &= u g g^{-1} \omega g + u g g^{-1} dg, \\ u' \omega' &= u' g^{-1} \omega g + u' g^{-1} dg, \end{aligned}$$

and so we obtain the **transformation law for connection forms**:

$$\omega' = g^{-1} \omega g + g^{-1} dg.$$

### 12.3. Differentiation Along a Map

Once again let  $\pi : E \rightarrow M$  be a vector bundle with a Koszul connection  $\nabla$ . Consider a smooth map  $f : N \rightarrow M$  and a section  $\sigma : N \rightarrow E$  along  $f$ . Let  $e_1, \dots, e_k$  be a frame field defined over  $U \subset M$ . Since  $f$  is continuous,  $O = f^{-1}(U)$  is open and  $e_1 \circ f, \dots, e_k \circ f$  are fields along  $f$  defined on  $O$ . We may write  $\sigma = \sum_{a=1}^k \sigma^a e_a \circ f$  for some functions  $\sigma^a : O \subset N \rightarrow \mathbb{F}$ .

For any  $p \in O$  and  $v \in T_p N$ , we define

$$(12.5) \quad \nabla_v^f \sigma := \sum_{a=1}^k \left( (d\sigma^a \cdot v) + \sum_{r=1}^k \omega_r^a(Tf(v)) \sigma^r(p) \right) e_a(f(p)).$$