5
Tensors

In the previous chapter we studied vector fields and functions on manifolds. In this chapter these objects are generalized to tensor fields, which are sections of vector bundles built out of the tangent bundle. This study is continued in the next chapter when we discuss differential forms, which are tensors with special symmetry properties. One of the objectives of this chapter is to extend the pull-back and Lie derivative operations from functions and vector fields to tensor fields.

5.1 Tensors on Linear Spaces

Preparatory to putting tensors on manifolds, we first study them on vector spaces. This subject is an extension of linear algebra sometimes called "multilinear algebra." Ultimately our constructions will be done on each fiber of the tangent bundle, producing a new vector bundle.

As in Chapter 2, $\mathbf{E}, \mathbf{F}, \ldots$ denote Banach spaces and $L^k(\mathbf{E}_1, \ldots, \mathbf{E}_k; \mathbf{F})$ denotes the vector space of continuous k-multilinear maps of $\mathbf{E}_1 \times \cdots \times \mathbf{E}_k$ to $\mathbf{F}$. The special case $L(\mathbf{E}, \mathbb{R})$ is denoted $\mathbf{E}^*$, the ***dual space*** of $\mathbf{E}$. If $\mathbf{E}$ is finite dimensional and $\{e_1, \ldots, e_n\}$ is an ordered basis of $\mathbf{E}$, there is a unique ordered basis of $\mathbf{E}^*$, the ***dual basis*** $\{e^1, \ldots, e^n\}$, such that $\langle e^j, e_i \rangle = \delta_i^j$ where $\delta_i^j = 1$ if $j = i$ and 0 otherwise. Furthermore, for each $v \in \mathbf{E}$,

$$v = \sum_{i=1}^{n} \langle e^i, v \rangle \, e_i \quad \text{and} \quad \alpha = \sum_{i=1}^{n} \langle \alpha, e_i \rangle \, e^i,$$

for each $\alpha \in \mathbf{E}^*$, where $\langle \, , \rangle$ denotes the pairing between $\mathbf{E}$ and $\mathbf{E}^*$. Employing the ***summation convention*** whereby summation is implied when an index is repeated on upper and lower levels, these expressions become

$$v = \langle e^i, v \rangle \, e_i \quad \text{and} \quad \alpha = \langle \alpha, e_i \rangle \, e^i.$$

As in Supplement 2.4C, if $\mathbf{E}$ is infinite dimensional, by $\mathbf{E}^*$ we will mean another Banach space weakly paired to $\mathbf{E}$; it need not be the full functional analytic dual of $\mathbf{E}$. In particular, $\mathbf{E}^{**}$ will *always* be chosen to be $\mathbf{E}$. With these conventions, tensors are defined as follows.

5.1.1 Definition. *For a vector space* $\mathbf{E}$ *we put*

$$T_s^r(\mathbf{E}) = L^{r+s}(\mathbf{E}^*, \ldots, \mathbf{E}^*, \mathbf{E}, \ldots, \mathbf{E}; \mathbb{R})$$

(r copies of $\mathbf{E}^*$ and s copies of $\mathbf{E}$). Elements of $T^r_s(\mathbf{E})$ are called ***tensors on $\mathbf{E}$, contravariant of order*** *r and* ***covariant of order*** *s; or simply, of* ***type*** *(r, s).*

Given $t_1 \in T^{r_1}_{s_1}(\mathbf{E})$ and $t_2 \in T^{r_2}_{s_2}(e)$, the ***tensor product*** *of t_1 and t_2 is the tensor $t_1 \otimes t_2 \in T^{r_1+r_2}_{s_1+s_2}(\mathbf{E})$ defined by*

$$\begin{aligned}(t_1 \otimes t_2)(\beta^1, \ldots, \beta^{r_1}, \gamma^1, \ldots, \gamma^{r_2}, f_1, \ldots, f_{s_1}, g_1, \ldots, g_{s_2})\\ = t_1(\beta^1, \ldots, \beta^{r_1}, f_1, \ldots, f_{s_1}) t_2(\gamma^1, \ldots, \gamma^{r_2}, g_1, \ldots, g_{s_2})\end{aligned}$$

where $\beta^j, \gamma^j \in \mathbf{E}^*$ and $f_j, g_j \in \mathbf{E}$.

Replacing $\mathbb{R}$ by a space $\mathbf{F}$ gives $T^r_s(\mathbf{E}; \mathbf{F})$, the $\mathbf{F}$-valued tensors of type (r, s). The tensor product now requires a bilinear form on the value space for its definition. For $\mathbb{R}$-valued tensors, $\otimes$ is associative, bilinear and continuous; it is *not* commutative. We also have the special cases

$$T^1_0(\mathbf{E}) = \mathbf{E},\ T^0_1(\mathbf{E}) = \mathbf{E}^*,\ T^0_2(\mathbf{E}) = L(\mathbf{E}; \mathbf{E}^*),\ \text{and } T^1_1(\mathbf{E}) = L(\mathbf{E}; \mathbf{E})$$

and make the convention that $T^0_0(\mathbf{E}; \mathbf{F}) = \mathbf{F}$.

5.1.2 Proposition. *Let $\mathbf{E}$ be an n dimensional vector space. If $\{e_1, \ldots, e_n\}$ is a basis of $\mathbf{E}$ and $\{e^1, \ldots, e^n\}$ is the dual basis, then*

$$\{\, e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s} \mid i_1, \ldots, i_r, j_1, \ldots, j_s = 1, \ldots, n \,\}$$

is a basis of $T^r_s(\mathbf{E})$ and thus $\dim(T^r_s(\mathbf{E})) = n^{r+s}$.

Proof. We must show that the elements

$$e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}$$

of $T^r_s(\mathbf{E})$ are linearly independent and span $T^r_s(\mathbf{E})$. Suppose that a finite sum vanishes:

$$t^{i_1 \ldots i_r}_{j_1 \ldots j_s} e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s} = 0.$$

Apply this to $(e^{k_1}, \ldots, e^{k_r}, e_{\ell_1}, \ldots, e_{\ell_s})$ to get $t^{k_1 \ldots k_r}_{\ell_1 \ldots \ell_s} = 0$. Next, check that for $t \in T^r_s(\mathbf{E})$ we have

$$t = t(e^{i_1}, \ldots, e^{i_r}, e_{j_1}, \ldots, e_{j_s}) e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}. \qquad \blacksquare$$

The coefficients

$$t^{i_1 \ldots i_r}_{j_1 \ldots j_s} = t(e^{i_1}, \ldots, e^{i_r}, e_{j_1}, \ldots, e_{j_s})$$

and called the ***components of t relative to the basis*** $\{e_1, \ldots, e_n\}$.

5.1.3 Examples.

A. If t is a $(0, 2)$-tensor on $\mathbf{E}$ then t has components $t_{ij} = t(e_i, e_j)$, an $n \times n$ matrix. This is the usual way of associating a bilinear form with a matrix. For instance, in $\mathbb{R}^2$ the bilinear form

$$t(x, y) = Ax_1y_1 + Bx_1y_2 + Cx_2y_1 + Dx_2y_2$$

(where $x = (x_1, x_2)$ and $y = (y_1, y_2)$) is associated to the 2×2 matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

B. If t is a $(0,2)$-tensor on $\mathbf{E}$, it makes sense to say that t is ***symmetric***; that is, $t(e_i, e_j) = t(e_j, e_i)$. This is equivalent to saying that the matrix $[t_{ij}]$ is symmetric. Symmetric $(0,2)$–tensors t can be recovered from their quadratic form

$$Q(e) = t(e, e) \quad \text{by} \quad t(e_1, e_2) = \frac{1}{4}\left[Q(e_1 + e_2) - Q(e_1 - e_2)\right],$$

the ***polarization identity***. If $\mathbf{E} = \mathbb{R}^2$ and t has the matrix

$$\begin{bmatrix} A & B \\ B & C \end{bmatrix}.$$

then $Q(x) = Ax_1^2 + 2Bx_1x_2 + Cx_2^2$. Symmetric $(0,2)$-tensors are thus closely related to quadratic forms and arise, for example, in mechanics as moment of inertia tensors and stress tensors.

C. In general, a ***symmetric*** $(r,0)$-***tensor*** is defined by the condition

$$t(\alpha^1, \ldots, \alpha^r) = t(\alpha^{\sigma(1)}, \ldots, \alpha^{\sigma(r)})$$

for all permutations σ of $\{1, \ldots, r\}$, and all elements $\alpha^1, \ldots, \alpha^r \in \mathbf{E}^*$. One may associate to t a homogeneous polynomial of degree r, $P(\alpha) = t(\alpha, \ldots, \alpha)$ and as in the case $r = 2$, P and t determine each other. A similar definition holds for $(0,s)$-tensors. It is clear that *a tensor is symmetric iff all its components in an arbitrary basis are symmetric.*

D. An inner product $\langle\,,\rangle$ on $\mathbf{E}$ is a symmetric $(0,2)$-tensor. Its matrix has components $g_{ij} = \langle e_i, e_j\rangle$. Thus g_{ij} is symmetric and positive definite. The components of the inverse matrix are written g^{ij}.

E. The space $L^k(\mathbf{E}_1, \ldots, \mathbf{E}_k; \mathbf{F})$ is isometric to $L^k(\mathbf{E}_{\sigma(1)}, \ldots, \mathbf{E}_{\sigma(k)}; \mathbf{F})$ for any permutation σ of $\{1, \ldots, k\}$, the isometry being given by $A \mapsto A'$, where

$$A'(e_{\sigma(1)}, \ldots, e_{\sigma(k)}) = A(e_1, \ldots, e_k).$$

Thus if $t \in T^r_s(\mathbf{E}; \mathbf{F})$, the tensor t can be regarded in $C(r+s, s)$ (the number of ways $r+s$ objects chosen s at a time) ways as an $(r+s)$-multilinear $\mathbf{F}$-valued map. For example, if $t \in T^2_1(\mathbf{E})$, the standard way is to regard it as a 3-linear map $t : \mathbf{E}^* \times \mathbf{E}^* \times \mathbf{E} \to \mathbb{R}$. There are two more ways to interpret this map, however, namely as $\mathbf{E}^* \times \mathbf{E} \times \mathbf{E}^* \to \mathbb{R}$ and as $\mathbf{E} \times \mathbf{E}^* \times \mathbf{E}^* \to \mathbb{R}$. In finite dimensions, where one writes the tensors in components, this distinction is important and is reflected in the index positions. Thus the three *different* tensors described above are written

$$t^{ij}_k e_i \otimes e_j \otimes e^k, \quad t^i{}_{j}{}_k e_i \otimes e^k \otimes e_j, \quad t^{ij}_k e^k \otimes e_i \otimes e_j.$$

F. In classical mechanics one encounters the notion of a ***dyadic*** (cf. Goldstein [1980]). A dyadic is the formal sum of a finite number of ***dyads***, a dyad being a pair of vectors $e_1, e_2 \in \mathbb{R}^3$ written in a specific order in the form e_1e_2. The action of a dyad on a pair of vectors, called the ***double dot product*** of two dyads is defined by

$$e_1e_2 : u_1u_2 = (e_1 \cdot u_1)(e_2 \cdot u_2),$$

where $\cdot$ stands for the usual dot product in $\mathbb{R}^3$. In this way dyads and dyadics are nothing but $(0,2)$-tensors on $\mathbb{R}^3$; that is, $e_1e_2 = e_1 \otimes e_2 \in T^0_2(\mathbb{R}^3)$, by identifying $(\mathbb{R}^3)^*$ with $\mathbb{R}^3$.

G. Higher order tensors arise in elasticity and Riemannian geometry. In elasticity, the stress tensor is a symmetric 2-tensor and the elasticity tensor is a fourth-order tensor (see Marsden and Hughes [1983]). In Riemannian geometry the metric tensor is a symmetric 2-tensor and the curvature tensor is a fourth-order tensor. ♦

Interior Product. The ***interior product*** of a vector $v \in \mathbf{E}$ (resp., a form $\beta \in \mathbf{E}^*$) with a tensor $t \in T^r_s(\mathbf{E};\mathbf{F})$ is the $(r, s-1)$ (resp., $(r-1, s)$) type $\mathbf{F}$-valued tensor defined by

$$(\mathbf{i}_v t)(\beta^1, \dots, \beta^r, v_1, \dots, v_{s-1}) = t(\beta^1, \dots, \beta^r, v, v_1, \dots, v_{s-1})$$
$$(\mathbf{i}^\beta t)(\beta^1, \dots, \beta^{r-1}, v_1, \dots, v_s) = t(\beta, \beta^1, \dots, \beta^{r-1}, v_1, \dots, v_s).$$

Clearly, $\mathbf{i}_v : T^r_s(\mathbf{E};\mathbf{F}) \to T^r_{s-1}(\mathbf{E};\mathbf{F})$ and $\mathbf{i}^\beta : T^r_s(\mathbf{E};\mathbf{F}) \to T^{r-1}_s(\mathbf{E};\mathbf{F})$ are linear continuous maps, as are $v \mapsto \mathbf{i}_v$ and $\beta \mapsto \mathbf{i}^\beta$. If $\mathbf{F} = \mathbb{R}$ and $\dim(\mathbf{E}) = n$, these operations take the following form in components. If e_k (resp., e^k) denotes the kth basis (resp., dual basis) element of $\mathbf{E}$, we have

$$\mathbf{i}_{e_k}(e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}) = \delta^{j_1}_k e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_2} \otimes \cdots \otimes e^{j_s},$$
$$\mathbf{i}^{e^k}(e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}) = \delta^k_{i_1} e_{i_s} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}.$$

By Proposition 5.1.2 these formulas and linearity enable us to compute any interior product.

Contractions. Let $\dim(\mathbf{E}) = n$. The contraction of the kth contravariant with the ℓth covariant index, or for short, the (k, ℓ)-***contraction***, is the family of linear maps $C^k_\ell : T^r_s(\mathbf{E};\mathbf{F}) \to T^{r-1}_{s-1}(\mathbf{E})$ defined for any pair of natural numbers $r, s \ge 1$ by

$$\begin{aligned} &C^k_\ell(T^{i_1\dots i_s}_{j_1\dots j_r} e_{i_1} \otimes \cdots \otimes e_{i_s} \otimes e^{j_1} \otimes \cdots \otimes e^{j_r}) \\ &\quad = t^{i_1\dots i_{k-1} p i_{k+1}\dots i_s}_{j_1\dots j_{\ell-1} p j_{\ell+1}\dots j_r} e_{i_1} \otimes \cdots \otimes \hat{e}_{i_k} \otimes \cdots \otimes e_{i_s} \otimes e^{j_1} \otimes \cdots \otimes \hat{e}^{j_\ell} \otimes \cdots \otimes e^{j_r}, \end{aligned}$$

where $\{e_1, \dots, e_n\}$ is a basis of $\mathbf{E}$, $\{e_1, \dots, e_n\}$ is the dual basis in $\mathbf{E}^*$, and $\hat{}$ over a vector or covector means that it is omitted. It is straightforward to verify that C^k_ℓ so defined is independent of the basis. This is essentially the same computation that is needed to show that the trace of a linear transformation of $\mathbf{E}$ to itself is intrinsic.

If $\mathbf{E}$ is infinite dimensional, contraction is not defined for arbitrary tensors. One introduces the so-called ***contraction class tensors***, analogous to the trace class operators, defines contraction as above in terms of a Banach space basis and its dual, and shows that the contraction class condition implies that the definition is basis independent. We shall not dwell upon these technicalities, refer to Rudin [1973] for a brief discussion of trace class operators, and invite the reader to model the concept of contraction class along these lines. For example, if $\mathbf{E}^* = \mathbf{E} = \ell^2(\mathbb{R})$, and $e_i = e^i$ equals the sequence with 1 in the ith place and zero everywhere else, then

$$t = \sum_{n=0}^{\infty} 2^{-n} e_n \otimes e^n \in T^1_1(\mathbf{E}) \quad \text{and} \quad C^1_1(t) = \sum_{n=0}^{\infty} 2^{-n} = 2.$$

Kronecker Delta. The ***Kronecker delta*** is the tensor $\delta \in T^1_1(\mathbf{E})$ defined by $\delta(\alpha, e) = \langle \alpha, e \rangle$. If $\mathbf{E}$ is finite dimensional, δ corresponds to the identity $I \in L(\mathbf{E};\mathbf{E})$ under the canonical isomorphism $T^1_1(\mathbf{E}) \cong L(\mathbf{E};\mathbf{E})$. Relative to any basis, the components of δ are the usual Kronecker symbols δ^i_j, that is, $\delta = \delta^i_j e_i \otimes e^j$.

Associated Tensors. Suppose $\mathbf{E}$ is a finite-dimensional real inner product space with a basis $\{e_1, \dots, e_n\}$ and corresponding dual basis $\{e^1, \dots, e^n\}$ in $\mathbf{E}^*$. Using the inner product, with matrix denoted by $[g_{ij}]$, so $g_{ij} = \langle\!\langle e_i, e_j \rangle\!\rangle$, we get the isomorphism

$$^\flat : \mathbf{E} \to \mathbf{E}^* \text{ given by } x \mapsto \langle\!\langle x, \cdot \rangle\!\rangle, \quad \text{and its inverse } ^\sharp : \mathbf{E}^* \to \mathbf{E}.$$

The matrix of $^\flat$ is $[g_{ij}]$; that is,

$$(x^\flat)_i = g_{ij} x^j$$

and of $^\sharp$ is $[g^{ij}]$; that is,

$$(\alpha^\sharp)^i = g^{ij} \alpha_j,$$

where x^j and α_j are the components of e and α, respectively. We call $^\flat$ the ***index lowering operator*** and $^\sharp$ the ***index raising operator***.

These operators can be applied to tensors to produce new ones. For example if t is a tensor of type $(0,2)$ we can define an ***associated tensor*** t' of type $(1,1)$ by

$$t'(e,\alpha) = t(e,\alpha^\sharp).$$

The components are

$$(t')_i^j = g^{jk} t_{ik} \quad \text{(as usual, sum on } k\text{)}.$$

In the classical literature one writes t_i^j for $g^{jk}t_{ik}$, and this is indeed a convenient notation in calculations. However, contrary to the impression one may get from the classical theory of Cartesian tensors, t and t' are *different tensors*.

5.1.4 Examples.

Let $\mathbf{E}$ be a finite-dimensional real vector space with basis $\{e_1, \ldots, e_n\}$ and dual basis $\{e^1, \ldots, e^n\}$.

A. If $t \in T_1^2(\mathbf{E})$ and $x = x^i e_i$, then

$$\begin{aligned} \mathbf{i}_x t &= x^p \mathbf{i}_{e_p}(t_j^{k\ell} e_k \otimes e_\ell \otimes e^j) = x^p t_j^{k\ell} \mathbf{i}_{e_p}(e_k \otimes e_\ell \otimes e^j) \\ &= x^p t_j^{k\ell} \delta_p^j e_k \otimes e_\ell = x^p t_p^{k\ell} e_k \otimes e_\ell. \end{aligned}$$

Thus, the components of $\mathbf{i}_x t$ are $x^p t_p^{k\ell}$. The interior product of the same tensor with $\alpha = \alpha_p e^p$ takes the form

$$\mathbf{i}^\alpha t = \alpha_p t_j^{k\ell} \mathbf{i}^{e^p}(e_k \otimes e_\ell \otimes e^j) = \alpha_p t_j^{k\ell} \delta_k^p e_\ell \otimes e^j = \alpha_p t_j^{p\ell} e_\ell \otimes e^j.$$

B. If $t \in T_3^2(\mathbf{E})$, the $(2,1)$-contraction is given by

$$C_1^2(t_{k\ell m}^{ij} e_i \otimes e_j \otimes e^k \otimes e^\ell \otimes e^m) = t_{j\ell m}^{ij} e_i \otimes e^\ell \otimes e^m.$$

C. An important particular example of contraction is the ***trace*** of a $(1,1)$-tensor. Namely, if $t \in T_1^1(\mathbf{E})$, then $\operatorname{trace}(t) = C_1^1(t) = t_i^i$, where $t = t_j^i e_i \otimes e^j$.

D. The components of the tensor associated to g by raising the second index are $g^{jk} g_{ik} = g^{jk} g_{ki} = \delta_i^j$.

E. Let

$$t \in T_2^3(\mathbf{E}), \quad t = t_{\ell m}^{ijk} e_i \otimes e_j \otimes e_k \otimes e^\ell \otimes e^m.$$

Then t has quite a few associated tensors, depending on which index is lowered or raised. For example

$$\begin{aligned} t_\ell^{ijkm} &= g^{mp} t_{\ell p}^{ijk}, \\ t_{jk\ell}^{im} &= g_{ja} g_{kb} g^{mc} t_{\ell c}^{iab} \\ t_{ij}^{k\ell m} &= g_{ia} g_{jb} g^{\ell c} g^{md} t_{cd}^{abk} \\ t_{ikm}^{j\ell} &= g_{ia} g_{kb} g^{\ell c} t_{cm}^{ajb} \end{aligned}$$

and so on.

F. The positioning of the indices in the components of associated tensors is important. For example, if $t \in T^0{}_2(\mathbf{E})$, we saw earlier that $t_i^j = g^{jk} t_{ik}$. However, $t_i^j = g^{ij} t_{ki}$, which is in general *different* from t_i^j when t is not symmetric. For example, if $\mathbf{E} = \mathbb{R}^3$ with $g_{ij} = \delta_{ij}$ and the nine components of t in the standard basis are $t_{12} = 1$, $t_{21} = -1$, $t_{ij} = 0$ for all other pairs i, j, then $t_i^j = t_{ij}$, $t^j{}_i = t_{ji}$, so that $t_1{}^2 = t_{12} = 1$ while $t^2{}_1 = t_{21} = -1$.

G. The ***trace*** of a $(2,0)$-tensor is defined to be the trace of the associated $(1,1)$ tensor; that is, if $t = t^{ij} e_i \otimes e_j$, then

$$\operatorname{trace}(t) = t_i^i = g_{ik} t^{ik}.$$

The question naturally arises whether we get the same answer by lowering the first index instead of the second, that is, if we consider t_i^i. By symmetry of g_{ij} we have $t_i^i = g_{ki} t^{ik} = t_k^k$, so that the definition of the trace is independent of which index is lowered. Similarly, if

$$t \in T_2^0(\mathbf{E}), \quad \operatorname{trace}(t) = t_i^i = g^{ik} t_{ik} = t_k^k.$$

In particular $\operatorname{trace}(g) = g_i^i = g_{ik}^{ik} = \dim(\mathbf{E})$. ♦

The Dual of a Linear Transformation. If $\varphi \in L(\mathbf{E}, \mathbf{F})$, the ***transpose*** or ***dual*** of φ, denoted $\varphi^* \in L(\mathbf{F}^*, \mathbf{E}^*)$ is defined by $\langle \varphi^*(\beta), e \rangle = \langle \beta, \varphi(e) \rangle$, where $\beta \in \mathbf{F}^*$ and $e \in \mathbf{E}$.

Let us analyze the matrices of φ and φ^*. As customary in linear algebra, vectors in a given basis are represented by a column whose entries are the components of the vector. Let $\varphi \in L(\mathbf{E}, \mathbf{F})$ and let $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_m\}$ be ordered bases of $\mathbf{E}$ and $\mathbf{F}$ respectively. Put $\varphi(e_i) = A_i^a f_a$. (We use a different dummy index for the $\mathbf{F}$-index to avoid confusion.) This defines the matrix of φ; $\mathbf{A} = [A_i^a]$. Thus, for $v = v^i e_i \in \mathbf{E}$ the components of $\varphi(v)$ are given by $\varphi(v)^a = A_i^a v^i$. Hence, thinking of v and $\varphi(v)$ as column vectors, this formula shows that $\varphi(v)$ *is computed by multiplying v on the* ***left*** *by* $\mathbf{A}$, *the matrix of* φ, as in elementary linear algebra. Thus, the upper index is the row index, while the lower index is the column index. Consequently, $\varphi(e_i)$ represents the ith column of the matrix of φ. Let us now investigate the matrix of $\varphi^* \in L(\mathbf{F}^*, \mathbf{E}^*)$. In the dual ordered bases, $\langle \varphi^*(f^a), e_i \rangle = \langle f^a, \varphi(e_i) \rangle = \left\langle f^a, A_i^b f_b \right\rangle = A_i^b \delta_b^a = A_i^a$, that is, $\varphi^*(f^a) = A_i^a e^i$ and thus $\varphi^*(f^a)$ is the ath row of $\mathbf{A}$. Consequently *the matrix of φ^* is the transpose of the matrix of φ.* If $\beta = \beta_a f^a \in \mathbf{F}^*$ then $\varphi^*(\beta) = \beta_a \varphi^*(f^a) = \beta_a A_i^a e^i$, which says that the ith component of $\varphi^*(\beta)$ equals $\varphi^*(\beta)_i = \beta_a A_i^a$. *Thinking of elements in the dual as rows whose entries are their components in the dual basis, this shows that $\varphi^*(\beta)$ is computed by multiplying β on the* ***right*** *by* $\mathbf{A}$, *the matrix of* φ, again in agreement with linear algebra.

Push-forward and Pull-back. Now we turn to the effect of linear transformations on tensors. We start with an induced map that acts "forward" like φ.

5.1.5 Definition. *If $\varphi \in L(\mathbf{E}, \mathbf{F})$ is an isomorphism, define the* ***push-forward*** *of φ, $T^r{}_s\varphi = \varphi_* \in L(T^r{}_s(\mathbf{E}), T^r{}_s(\mathbf{F}))$ by*

$$\varphi_* t(\beta^1, \ldots, \beta^r, f_1, \ldots, f_s) = t(\varphi^*(\beta^1), \ldots, \varphi^*(\beta^r), \varphi^{-1}(f_1), \ldots, \varphi^{-1}(f_s)),$$

where $t \in T^r{}_s(\mathbf{E})$, $\beta^1, \ldots, \beta^r \in \mathbf{F}^*$, and $f_1, \ldots, f_s \in \mathbf{F}$.

We leave the verification that φ_* is continuous to the reader. Note that $T^0{}_1\varphi = (\varphi^{-1})^*$, which maps "forward" like φ. If $\mathbf{E}$ and $\mathbf{F}$ are finite dimensional, then $T^1{}_0(\mathbf{E}) = \mathbf{E}$, $T^1{}_0(\mathbf{F}) = \mathbf{F}$ and we identify φ with $T^1{}_0\varphi$. The next proposition asserts that the push-forward operation is compatible with compositions and the tensor product.

5.1.6 Proposition. *Let $\varphi : \mathbf{E} \to \mathbf{F}$ and $\psi : \mathbf{F} \to \mathbf{G}$ be isomorphisms. Then*

(i) $(\psi \circ \varphi)_* = \psi_* \circ \varphi_*$;

(ii) *if $i : \mathbf{E} \to \mathbf{E}$ is the identity, then so is $i_* : T^r{}_s(\mathbf{E}) \to T^r{}_s(\mathbf{E})$;*

(iii) $\varphi_* : T^r{}_s(\mathbf{E}) \to T^r{}_s(\mathbf{F})$ *is an isomorphism, and* $(\varphi_*)^{-1} = (\varphi^{-1})_*$;

(iv) *If* $t_1 \in T^{r_1}{}_{s_1}(\mathbf{E})$ *and* $t_2 \in T^{r_2}{}_{s_1}(\mathbf{E})$, *then* $\varphi_*(t_1 \otimes t_2) = \varphi_*(t_1) \otimes \varphi_*(t_2)$.

Proof. For (i),

$$\begin{aligned}
&\psi_*(\varphi_* t)(\gamma^1, \ldots, \gamma^r, g_1, \ldots, g_s)\\
&\quad = \varphi_* t(\psi^*(\gamma^1), \ldots, \psi^*(\gamma^r), \psi^{-1}(g_1), \ldots, \psi^{-1}(g_s))\\
&\quad = t(\varphi^*\psi^*(\gamma^1), \ldots, \varphi^*\psi^*(\gamma^r), \varphi^{-1}\psi^{-1}(g_1), \ldots, \varphi^{-1}\psi^{-1}(g_s))\\
&\quad = t((\psi \circ \varphi)^*(\gamma^1), \ldots, (\psi \circ \varphi)^*(\gamma^r), (\psi \circ \varphi)^{-1}(g_1), \ldots, (\psi \circ \varphi)^{-1}(g_s))\\
&\quad = (\psi \circ \varphi)_* t(\gamma^1, \ldots, \gamma^r, g_1, \ldots, g_s),
\end{aligned}$$

where $\gamma^1, \ldots, \gamma^r \in \mathbf{G}^*$, $g_1, \ldots, g_s \in \mathbf{G}$, and $t \in T^r{}_s(\mathbf{E})$. We have used the fact that the transposes and inverses satisfy $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$ and $(\psi \circ \varphi)^{-1} = \varphi^{-1} \circ \psi^{-1}$, which the reader can easily check. Part (ii) is an immediate consequence of the definition and the fact that $i^* = i$ and $i^{-1} = i$. Finally, for (iii) we have $\varphi_* \circ (\varphi^{-1})_* = i_*$, the identity on $T^r{}_s(\mathbf{F})$, by (i) and (ii). Similarly, $(\varphi^{-1})_* \circ \varphi_* = i_*$ the identity on $T^r{}_s(\mathbf{E})$. Hence (iii) follows. Finally (iv) is a straightforward consequence of the definitions. ■

Since $(\varphi^{-1})_*$ maps "backward" it is called the ***pull-back*** of φ and is denoted φ^*. The next proposition gives a connection with component notation.

5.1.7 Proposition. *Let* $\varphi \in L(\mathbf{E}, \mathbf{F})$ *be an isomorphism of finite dimensional vector spaces. Let* $[A^a_i]$ *denote the matrix of* φ *in the ordered bases* $\{e_1, \ldots, e_n\}$ *of* $\mathbf{E}$ *and* $\{f_1, \ldots, f_n\}$ *of* $\mathbf{F}$, *that is,* $\varphi(e_i) = A^a_i f_a$. *Denote by* $[B^i_a]$ *the matrix of* φ^{-1}, *that is,* $\varphi^{-1}(f_a) = B^i_a e_i$. *Then* $[B^i_a]$ *is the inverse matrix of* $[A^a_i]$ *in the sense that* $B^i_a A^a_j = \delta^i_j$. *Let*

$$t \in T^r{}_s(\mathbf{E}) \quad \textit{with components} \quad t^{i_1 \ldots i_r}_{j_1 \ldots j_s} \quad \textit{relative to} \quad \{e_1, \ldots, e_n\}$$

and

$$q \in T^r{}_s(\mathbf{F}) \quad \textit{with components} \quad q^{a_1 \ldots a_r}_{b_1 \ldots b_s} \quad \textit{relative to} \quad \{f_1, \ldots, f_n\}.$$

Then the components of $\varphi_* t$ *relative to* $\{f_1, \ldots, f_n\}$ *and of* $\varphi^* q$ *relative to* $\{e_1, \ldots, e_n\}$ *are given respectively by*

$$\begin{aligned}
(\varphi_* t)^{a_1 \ldots a_r}_{b_1 \ldots b_s} &= A^{a_1}_{i_1} \ldots A^{a_r}_{i_r} t^{i_1 \ldots i_r}_{j_1 \ldots j_s} B^{j_1}_{b_1} \ldots B^{j_s}_{b_s}\\
(\varphi_* q)^{i_1 \ldots i_r}_{j_1 \ldots j_s} &= B^{i_1}_{a_1} \ldots B^{i_r}_{a_r} q^{a_1 \ldots a_r}_{b_1 \ldots b_s} A^{b_1}_{j_1} \ldots A^{b_s}_{j_s}.
\end{aligned}$$

Proof. We have

$$e_i = \varphi^{-1}(\varphi(e_i)) = \varphi^{-1}(A^a_i f_a) = A^a_i \varphi^{-1}(f_a) = A^a_i B^j_a e_j,$$

whence $B^j_a A^a_i = \delta^j_i$ for all i, j. Similarly, one shows that $A^b_i B^i_a = \delta^b_a$, so that $[A^a_i]^{-1} = [B^j_a]$. We have

$$\begin{aligned}
(\varphi_* t)^{a_1 \ldots a_r}_{b_1 \ldots b_s} &= (\varphi_* t)(f^{a_1}, \ldots, f^{a_r}, f_{b_1}, \ldots, f_{b_s})\\
&= t(\varphi^*(f^{a_1}), \ldots, \varphi^*(f^{a_r}), \varphi^{-1}(f_{b_1}), \varphi^{-1}(f_{b_s}))\\
&= t(A^{a_1}_{i_1} e^{i_1}, \ldots, A^{a_r}_{i_r} e^{i_r}, B^{j_1}_{b_1} e_{j_1}, \ldots, B^{j_s}_{b_s} e_{i_s})\\
&= A^{a_1}_{i_1} \ldots A^{a_r}_{i_r} t^{i_1 \ldots i_r}_{j_1 \ldots j_s} b^{j_1}_{b_1} \ldots b^{j_s}_{b_s}.
\end{aligned}$$

To prove the second relation, we need the matrix of $(\varphi^{-1})^* \in L(\mathbf{E}^*, \mathbf{F}^*)$. We have

$$\left\langle (\varphi^{-1})^*(e^i), f_a \right\rangle = \left\langle e^i, \varphi^{-1}(f_a) \right\rangle = \left\langle e^i, B^k_a e_k \right\rangle = B^i_a$$

so that $(\varphi^{-1})^*(e^i) = B^i_a f^a$. Now proceed as in the previous case. ■

Note that *the matrix of* $(\varphi^{-1})^* \in L(\mathbf{E}^*, \mathbf{F}^*)$ *is the transpose of the inverse of the matrix of* φ.

The assumption that φ be an isomorphism for φ_* to exist is quite restrictive but clearly cannot be weakened. However, one might ask if instead of "push-forward," the "pull-back" operation is considered, this restrictive assumption can be dropped. This is possible when working with covariant tensors, even when $\varphi \in L(\mathbf{E}, \mathbf{F})$ is arbitrary.

5.1.8 Definition. *If* $\varphi \in L(\mathbf{E}, \mathbf{F})$ *(not necessarily an isomorphism), define the* ***pull-back*** $\varphi^* \in L({T^0}_s(\mathbf{F}), {T^0}_s(\mathbf{E}))$ *by*

$$\varphi^* t(e_1, \ldots, e_s) = t(\varphi(e_1), \ldots, \varphi(e_s)),$$

where $t \in {T^0}_s(\mathbf{F})$ *and* $e_1, \ldots, e_s \in \mathbf{E}$.

Likewise, one can push forward tensors in $T^r_0(\mathbf{E})$ even if φ is not an isomorphism.

The next proposition asserts that φ^* is compatible with compositions and the tensor product. Its proof is almost identical to that of proposition 5.1.6 and is left as an exercise for the reader.

5.1.9 Proposition. *Let* $\varphi \in L(\mathbf{E}; \mathbf{F})$ *and* $\psi \in L(\mathbf{F}; \mathbf{G})$.

(i) $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$.

(ii) *If* $i : \mathbf{E} \to \mathbf{E}$ *is the identity, then so is* $i^* \in L(T^0_s(\mathbf{E}), T^0_s(\mathbf{E}))$.

(iii) *If* φ *is an isomorphism, then so is* φ^* *and*

$$\varphi^* = (\varphi^{-1})_*.$$

(iv) *If* $t_1 \in T^0_{s_1}(\mathbf{F})$ *and* $t_2 \in T^0_{s_2}(\mathbf{F})$, *then*

$$\varphi^*(t_1 \otimes t_2) = (\varphi^* t_1) \otimes (\varphi^* t_2).$$

Finally, the components of $\varphi^* t$ are given by the following.

5.1.10 Proposition. *Let* $\mathbf{E}$ *and* $\mathbf{F}$ *be finite-dimensional vector spaces and* $\varphi \in L(\mathbf{E}, \mathbf{F})$. *For ordered bases* $\{e_1, \ldots, e_n\}$ *of* $\mathbf{E}$ *and* $\{f_1, \ldots, f_m\}$ *of* $\mathbf{F}$, *suppose that* $\varphi(e_i) = A^a_i f_a$, *and let* $t \in T^0_s(\mathbf{F})$ *have components* $t_{b_1 \ldots b_s}$. *Then the components of* $\varphi^* t$ *relative to* $\{e_1, \ldots, e_n\}$ *are given by*

$$(\varphi^* t)_{j_1 \ldots j_s} = t_{b_1 \ldots b_s} A^{b_1}_{j_1} \ldots A^{b_s}_{j_s}.$$

Proof.

$$\begin{aligned}(\varphi^* t)_{j_1 \ldots j_s} &= (\varphi^* t)(e_{j_1}, \ldots, e_{j_s}) = t(\varphi(e_{j_1}), \ldots, \varphi(e_{j_s})) \\ &= t(A^{b_1}_{j_1} f_{b_1}, \ldots, A^{b_s}_{j_s} f_{b_s}) = t(f_{b_1}, \ldots, f_{b_s}) A^{b_1}_{j_1} \ldots A^{b_s}_{j_s} \\ &= t_{b_1 \ldots b_s} A^{b_1}_{j_1} \ldots A^{b_s}_{j_s}\end{aligned}$$

■

5.1.11 Examples.

A. On $\mathbb{R}^2$ with the standard basis $\{e_1, e_2\}$, let $t \in T^2_0(\mathbb{R}^2)$ be given by $t = e_1 \otimes e_1 + 2 e_1 \otimes e_2 - e_2 \otimes e_1 + 3 e_2 \otimes e_2$ and let $\varphi \in L(\mathbb{R}^2, \mathbb{R}^2)$ have the matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Then φ is clearly an isomorphism, since $\mathbf{A}$ has an inverse matrix given by

$$\mathbf{B} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}.$$

According to Proposition 5.1.7, the components of $T = \varphi^* t$ relative to the standard basis of $\mathbb{R}^2$ are given by

$$T^{ij} = B^i_a B^j_b t^{ab}, \quad \text{with} \quad B^1_1 = 1,\ B^2_1 = B^1_2 = -1, \text{ and } B^2_2 = 2$$

so that

$$\begin{aligned} T^{12} &= B^1_1 B^2_1 t^{11} + B^1_1 B^2_2 t^{12} + B^1_2 B^2_1 t^{21} + B^1_2 B^2_2 t^{22} \\ &= 1 \cdot (-1) \cdot 1 + 1 \cdot 2 \cdot 2 + (-1) \cdot (-1) \cdot (-1) + (-1) \cdot 2 \cdot 3 = -4, \end{aligned}$$

$$\begin{aligned} T^{21} &= B^2_1 B^1_1 t^{11} + B^2_1 B^1_2 t^{12} + B^2_2 B^1_1 t^{21} + B^2_2 B^1_2 t^{22} \\ &= (-1) \cdot 1 \cdot 1 + (-1) \cdot (-1) \cdot 2 + 2 \cdot 1 \cdot (-1) + 2 \cdot (-1) \cdot 3 = -7, \end{aligned}$$

$$\begin{aligned} T^{11} &= B^1_1 B^1_1 t^{11} + B^1_1 B^1_2 t^{12} + B^1_2 B^1_1 t^{21} + B^1_2 B^1_2 t^{22} \\ &= 1 \cdot 1 \cdot 1 + 1 \cdot (-1) \cdot 2 + (-1) \cdot 1 \cdot (-1) + (-1) \cdot (-1) \cdot 3 = 3, \end{aligned}$$

$$\begin{aligned} T^{22} &= B^2_1 B^2_1 t^{11} + B^2_1 B^2_2 t^{12} + B^2_2 B^2_1 t^{21} + B^2_2 B^2_2 t^{22} \\ &= (-1) \cdot (-1) \cdot 1 + (-1) \cdot 2 \cdot 2 + 2 \cdot (-1) \cdot (-1) + 2 \cdot 2 \cdot 3 = 11. \end{aligned}$$

Thus, $\varphi^* t = 3e_1 \otimes e_1 - 4e_1 \otimes e_2 - 7e_2 \otimes e_1 + 11e_2 \otimes e_2$.

B. Let $t = e_1 \otimes e^2 - 2e_2 \otimes e^2 \in T^1_1(\mathbb{R}^2)$ and consider the same map $\varphi \in L(\mathbb{R}^2, \mathbb{R}^2)$ as in part (a) above. We could compute the components of $\varphi_* t$ relative to the standard basis of $\mathbb{R}^2$ using the formula in Proposition 5.1.7 as before. An alternative way to proceed directly using Proposition 5.1.6(iv), that is, the fact that φ_* is compatible with tensor products. Thus

$$\varphi_* t = \varphi_*(e_1 \otimes e^2 - 2e_2 \otimes e^2) = \varphi(e_1) \otimes \varphi_*(e^2) - 2\varphi(e_2) \otimes \varphi_*(e^2).$$

But $\varphi(e_1) = 2e_1 + e_2$, $\varphi(e_2) = e_1 + e_2$, and $\varphi_*(e^2) = -e^1 + 2e^2$, so that

$$\begin{aligned} \varphi_* t &= (2e_1 + e_2) \otimes (-e^1 + 2e^2) - 2(e_1 + e_2) \otimes (-e^1 + 2e^2) \\ &= e_2 \otimes e^1 - 2e_2 \otimes e^2. \end{aligned}$$

C. Let e_1, e_2 be the standard basis of $\mathbb{R}^2$ and e^1, e^2 be the dual basis, as usual. Let $t = -2e^1 \otimes e^2 \in T^0_2(\mathbb{R}^2)$ and $\varphi \in L(\mathbb{R}^3, \mathbb{R}^2)$ be given by the matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \\ 0 & -1 & 1 \end{bmatrix}.$$

We will compute $\varphi^* t \in T^0_2(\mathbb{R}^3)$ by using the fact that φ^* is compatible with tensor products and that the matrix of $\varphi^* \in L(\mathbb{R}^2, \mathbb{R}^3)$ is the transpose of $\mathbf{A}$. Recall that $\varphi^*(e^i)$ is the ith *row*, since matrices act on the right on covectors. Let f^1, f^2, f^3 denote the standard dual basis of $\mathbb{R}^3$. Then $\varphi^*(e^1) = f^1 + 2f^3$ and $\varphi^*(e^2) = -f^2 + f^3$, so that

$$\begin{aligned} \varphi^*(t) &= -2\varphi^*(e^1) \otimes \varphi^*(e^2) = -2(f^1 + 2f^3) \otimes (-f^2 + f^3) \\ &= 2f^1 \otimes f^2 - 2f^1 \otimes f^3 + 4f^3 \otimes f^2 - 4f^3 \otimes f^3. \end{aligned}$$

♦

Exercises

◇ **5.1-1.** Compute the interior product of the tensor

$$t = e_1 \otimes e_1 \otimes e^2 + 3e_2 \otimes e_2 \otimes e^1$$

with $e = -e_1 + 2e_2$ and $\alpha = 2e^1 + e^2$. What are the $(1,1)$ and $(2,1)$ contractions of t?

◇ **5.1-2.** Compute all associated tensors of $t = e_1 \otimes e^2 \otimes e^2 + 2e_2 \otimes e^1 \otimes e^2 - e_2 \otimes e^2 \otimes e^1$ with respect to the standard metric of $\mathbb{R}^2$.

◇ **5.1-3.** Let $t = 2e^1 \otimes e^1 - e^2 \otimes e^1 + 3e^1 \otimes e^2$ and $\varphi \in L(\mathbb{R}^2, \mathbb{R}^2)$, $\psi \in L(\mathbb{R}^3, \mathbb{R}^2)$ be given by the matrices

$$\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 2 \end{bmatrix}.$$

Compute: trace(t), $\varphi^* t$, $\psi^* t$, trace$(\varphi^* t)$, trace$(\psi^* t)$, $\varphi_* t$, and all associated tensors of t, $\varphi^* t$, $\psi^* t$, and $\varphi_* t$ with respect to the corresponding standard inner products in $\mathbb{R}^2$ and $\mathbb{R}^3$.

◇ **5.1-4.** Let $\dim(\mathbf{E}) = n$ and $\dim(\mathbf{F}) = m$. Show that $T^r_s(\mathbf{E}; \mathbf{F})$ is an mn^{r+s}-dimensional real vector space by exhibiting a basis.

5.2 Tensor Bundles and Tensor Fields

We now extend the tensor algebra to local vector bundles, and then to vector bundles. For $U \subset \mathbf{E}$ (open) recall that $U \times \mathbf{F}$ is a local vector bundle. Then $U \times T^r_s(\mathbf{F})$ is also a local vector bundle in view of Proposition 5.1.2. Suppose $\varphi : U \times \mathbf{F} \to U' \times \mathbf{F}'$ is a local vector bundle mapping and is an *isomorphism on each fiber*; that is $\varphi_u = \varphi | \{u\} \times \mathbf{F} \in L(\mathbf{F}, \mathbf{F}')$ is an isomorphism. Also, let φ_0 denote the restriction of φ to the zero section. Then φ induces a mapping of the local tensor bundles as follows.

5.2.1 Definition. *If $\varphi : U \times \mathbf{F} \to U' \times \mathbf{F}'$ is a local vector bundle mapping such that for each $u \in U$, φ_u is an isomorphism, let $\varphi_* : U \times T^r_s(\mathbf{F}) \to U' \times T^r_s(\mathbf{F}')$ be defined by*

$$\varphi_*(u, t) = (\varphi_0(u), (\varphi_u)_* t),$$

where $t \in T^r_s(\mathbf{F})$.

Before proceeding, we shall pause and recall some useful facts concerning linear isomorphisms from Lemmas 2.5.4 and 2.5.5.

5.2.2 Proposition. *Let* $\mathrm{GL}(\mathbf{E}, \mathbf{F})$ *denote the set of linear isomorphisms from* $\mathbf{E}$ *to* $\mathbf{F}$. *Then* $\mathrm{GL}(\mathbf{E}, \mathbf{F}) \subset L(\mathbf{E}, \mathbf{F})$ *is open.*

5.2.3 Proposition. *Define the maps*

$$\mathcal{A} : L(\mathbf{E}, \mathbf{F}) \to L(\mathbf{F}^*, \mathbf{E}^*); \quad \varphi \mapsto \varphi^*$$

and

$$\mathcal{I} : \mathrm{GL}(\mathbf{E}, \mathbf{F}) \to \mathrm{GL}(\mathbf{F}, \mathbf{E}); \quad \varphi \mapsto \varphi^{-1}.$$

Then $\mathcal{A}$ and $\mathcal{I}$ are of class C^∞ and

$$\mathbf{D}\mathcal{I}^{-1}(\varphi) \cdot \psi = -\varphi^{-1} \circ \psi \circ \varphi^{-1}.$$

Smoothness of $\mathcal{A}$ is clear since it is linear.

5.2.4 Proposition. *If $\varphi : U \times \mathbf{F} \to U' \times \mathbf{F}'$ is a local vector bundle map and φ_u is an isomorphism for all $u \in U$, then $\varphi_* : U \times T^r_s(\mathbf{F}) \to U' \times T^r_s(\mathbf{F}')$ is a local vector bundle map and $(\varphi_u)_* = (\varphi_*)_u$ is an isomorphism for all $u \in U$. Moreover, if φ is a local vector bundle isomorphism then so is φ_*.*

Proof. That φ_* is an isomorphism on fibers follows from Proposition 5.1.6(iii) and the last assertion follows from the former. By Definition 5.2.1 we need only establish that $(\varphi_u)_* = (\varphi_*)_u$ is of class C^∞. Now φ_u is a smooth function of u, and, by Proposition 5.2.3 φ_u^* and φ_u^{-1} are smooth functions of u. The map $(\varphi_u)_*$ is a Cartesian product of r factors φ_u^* and s factors φ_u^{-1}, so is smooth. Hence, from the product rule, $(\varphi_u)_*$ is smooth. ■

This smoothness can be verified also for finite-dimensional bundles by using the standard bases in the tensor spaces as local bundle charts and proving that the components of $\varphi_* t$ are C^∞ functions.

We have the following commutative diagram, which says that φ_* preserves fibers:

$$\begin{array}{ccc} U \times T^r_s(\mathbf{F}) & \xrightarrow{\ \varphi_*\ } & U' \times T^r_s(\mathbf{F}') \\ \Big\downarrow{\scriptstyle \pi} & & \Big\downarrow{\scriptstyle \pi'} \\ U & \xrightarrow[\ \varphi_0\ ]{} & U' \end{array}$$

Tensor Bundles. With the above preliminaries out of the way, we can now define tensor bundles.

5.2.5 Definition. *Let $\pi : E \to B$ be a vector bundle with $E_b = \pi^{-1}(b)$ denoting the fiber over the point $b \in B$. Define*

$$T^r_s(E) = \bigcup_{b \in B} T^r_s(E_b)$$

and $\pi^r_s : T^r_s(E) \to B$ by $\pi^r_s(e) = b$ where $e \in T^r_s(E_b)$. Furthermore, for a given subset A of B, we define

$$T^r_s(E)|A = \bigcup_{b \in A} T^r_s(E_b).$$

If $\pi' : E' \to B'$ is another vector bundle and $(\varphi, \varphi_0) : E \to E'$ is a vector bundle mapping with $\varphi_b = \varphi|E_b$ an isomorphism for all $b \in B$, let $\varphi_* : T^r_s(E) \to T^r_s(E')$ be defined by $\varphi_*|T^r_s(E_b) = (\varphi_b)_*$.

Now suppose that $(E|U, \varphi)$ is an admissible local bundle chart of π, where $U \subset B$ is an open set. Then the mapping $\varphi_*|[T^r_s(E)|U]$ is obviously a bijection onto a local bundle, and thus is a local bundle chart. Further, $(\varphi_*)_b = (\varphi_b)_*$ is a linear isomorphism, so this chart preserves the linear structure of each fiber. We shall call such a chart a ***natural chart*** of $T^r_s(E)$.

5.2.6 Theorem. *If $\pi : E \to B$ is a vector bundle, then the set of all natural charts of $\pi^r_s : T^r_s(E) \to B$ is a vector bundle atlas.*

Proof. Condition **VB1** is obvious. For **VB2**, suppose we have two overlapping natural charts, φ_* and ψ_*. For simplicity, let them have the same domain. Then $\alpha = \psi \circ \varphi^{-1}$ is a local vector bundle isomorphism, and by Proposition 5.1.6, $\psi_* \circ (\varphi_*)^{-1} = \alpha_*$, is a local vector bundle isomorphism by Proposition 5.2.4. ■

This atlas of natural charts called the ***natural atlas*** of π^r_s, generates a vector bundle structure, and it is easily seen that the resulting vector bundle is Hausdorff, and all fibers are isomorphic Banachable spaces. Hereafter, $T^r_s(E)$ will denote all of this structure.

5.2.7 Proposition. *If $f : E \to E'$ is a vector bundle map that is an isomorphism on each fiber, then $f_* : T^r_s(E) \to T^r_s(E')$ is also a vector bundle map that is an isomorphism on each fiber.*

Proof. Let (U, φ) be an admissible vector bundle chart of E, and let (V, ψ) be one of E' so that $f(U) \subset V$ and $f_{\varphi\psi} = \psi \circ f \circ \varphi^{-1}$ is a local vector bundle mapping. Then using the natural atlas, we see that

$$(f_*)_{\varphi_*, \psi_*} = (f_{\varphi\psi})_*. \qquad \blacksquare$$

5.2.8 Proposition. *Let $f : E \to E'$ and $g : E' \to E''$ be vector bundle maps that are isomorphisms on each fiber. Then so is $g \circ f$, and*

(i) $(g \circ f)_* = g_* \circ f_*$;

(ii) *if $i : E \to E$ is the identity, then $i_* : T^r_s(E) \to T^r_s(E)$ is the identity;*

(iii) *if $f : E \to E'$ is a vector bundle isomorphism, then so if f_* and $(f_*)^{-1} = (f^{-1})_*$.*

Proof. For (i) we examine representatives of $(g \circ f)_*$ and $g_* \circ f_*$. These representatives are the same in view of Proposition 5.1.6. Part (ii) is clear from the definition, and (iii) follows from (i) and (ii) by the same method as in Proposition 5.1.6. ■

We now specialize to the case where $\pi : E \to B$ is the tangent vector bundle of a manifold.

5.2.9 Definition. *Let M be a manifold and $\tau_M : TM \to M$ its tangent bundle. We call $T^r_s(M) = T^r_s(TM)$ the **vector bundle of tensors contravariant order** r **and covariant order** s, or simply of **type** (r, s). We identify $T^1_0(M)$ with TM and call $T^0_1(M)$ the **cotangent bundle** of M also denoted by $\tau^*_M : T^*M \to M$. The zero section of $T^r_s(M)$ is identified with M.*

Tensor Fields. Recall that a section of a vector bundle assigns to each base point b a vector in the fiber over b and the addition and scalar multiplication of sections takes place within each fiber. In the case of $T^r_s(M)$ these vectors are called ***tensors***. The C^∞ sections of $\pi : E \to B$ were denoted $\Gamma^\infty(\pi)$, or $\Gamma^\infty(E)$. Recall that $\mathcal{F}(M)$ denotes the set of mappings from M into $\mathbb{R}$ that are of class C^∞ (the standard local manifold structure being used on $\mathbb{R}$) together with its structure as a ring; namely, $f + g$, cf, fg for $f, g \in \mathcal{F}(M)$, $c \in \mathbb{R}$ are defined by

$$(f + g)(x) = f(x) + g(x), \quad (cf)(x) = c(f(x)), \quad \text{and} \quad (fg)(x) = f(x)g(x).$$

Finally, recall that a ***vector field*** on M is an element of $\mathfrak{X}(M) = \Gamma^\infty(TM)$.

5.2.10 Definition. *A **tensor field of type** (r, s) on a manifold M is a C^∞ section of $T^r_s(M)$. We denote by $\mathcal{T}^r_s(M)$ the set $\Gamma^\infty(T^r_s(M))$ together with its (infinite-dimensional) real vector space structure. A **covector field** or **differential one-form** is an element of $\mathfrak{X}^*(M) = \mathcal{T}^0_1(M)$.*

If $f \in \mathcal{F}(M)$ and $t \in \mathcal{T}^r_s(M)$, let $ft : M \to T^r_s(M)$ be defined by $m \mapsto f(m)t(m)$. If $X_i \in \mathfrak{X}(M)$, $i = 1, \ldots, s$, $\alpha^j \in \mathfrak{X}^*(M)$, $j = 1, \ldots, r$, and $t' \in \mathcal{T}^{r'}_{s'}(M)$ define

$$t(\alpha^1, \ldots, \alpha^r, X_1, \ldots, X_s) : M \to \mathbb{R} \quad \text{by} \quad m \mapsto t(m)(\alpha_1(m), \ldots, X_s(m))$$

and

$$t \otimes t' : M \to T^{r+r'}_{s+s'}(M) \quad \text{by} \quad m \mapsto t(m) \otimes t'(m).$$

5.2.11 Proposition. *With f, t, X_i, α^j, and t' as in Definition 5.2.10,*

$$ft \in \mathcal{T}^r_s(M), \quad t(\alpha^1, \ldots, X_s) \in \mathcal{F}(M), \quad \text{and} \quad t \otimes t' \in \mathcal{T}^{r+r'}_{s+s'}.$$

Proof. The differentiability is evident in each case from the product rule in local representation. ■

For the tangent bundle TM, a natural chart is obtained by taking $T\varphi$, where φ is an admissible chart of M. This in turn induces a chart $(T\varphi)_*$ on $T^r_s M$. We shall call these the ***natural charts*** of $T^r_s M$.

Coordinate Representation of Tensor Fields. Recall that $\partial/\partial x^i = (T\varphi)^{-1}(e_i)$, for $\varphi : U \to U' \subset \mathbb{R}^n$ a chart on M, is a basis of $\mathfrak{X}(U)$. The vector field $\partial/\partial x^i$ corresponds to the derivation $f \mapsto \partial f/\partial x^i$. Since $dx^i(\partial/\partial x^j) = \partial x^i/\partial x^j = \delta^i_j$, we see that dx^i is the dual basis of $\partial/\partial x^i$ at every point of U, that is, that $dx^i = \varphi^*(e^i)$, where $\{e^1, \ldots, e^n\}$ is the dual basis to $\{e_1, \ldots, e_n\}$. Let

$$t^{i_1\ldots i_r}_{j_1\ldots j_s} = t\left(dx^{i_1}, \ldots, dx^{i_r}, \frac{\partial}{\partial x^{j_1}}, \ldots, \frac{\partial}{\partial x^{j_s}}\right) \in \mathcal{F}(U).$$

Applying Proposition 5.1.6(iv) at every point yields the coordinate expression of an (r,s)-tensor field:

$$t\,\big|U = t^{i_1\ldots i_r}_{j_1\ldots j_s}\,\frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}.$$

To discuss the behavior of these components relative to a change of coordinates, assume that $X^i : U \to \mathbb{R}$, $i = 1, \ldots, n$ is a different coordinate system. We can write $\partial/\partial x^i = a^j_i \partial/\partial X^j$, since both are bases of $\mathfrak{X}(U)$. Applying both sides to X^k yields $a^j_i = \partial X^j/\partial x^i$ that is, $(\partial/\partial x^i) = (\partial X^j/\partial x^i)(\partial/\partial X^j)$. Thus the dx^i, as dual basis change with the inverse of the Jacobian matrix $[\partial X^j/\partial x^i]$; that is, $dx^i = (\partial x^i/\partial X^j)dX^j$. Writing t in both coordinate systems and isolating equal terms gives the following change of coordinate formula for the components:

$$T^{k_1\ldots k_r}_{\ell_1\ldots \ell_s} = \frac{\partial X^{k_1}}{\partial x^{i_1}} \cdots \frac{\partial X^{k_r}}{\partial x^{i_r}} \frac{\partial x^{j_1}}{\partial X^{\ell_1}} \cdots \frac{\partial x^{j_s}}{\partial X^{\ell_s}} t^{i_1\ldots i_r}_{j_1\ldots j_s}$$

This formula is known as the ***tensoriality criterion***: *A set of n^{r+s} functions $t^{i_1\ldots i_r}_{j_1\ldots j_s}$ defined for each coordinate system on the open set U of M locally define an (r,s)-tensor field iff changes of coordinates have the aforementioned effect on them.* This statement is clear since at every point it assures that the n^{r+s} functions are the components of an (r,s)-tensor in T_uU and conversely.

The algebraic operations on tensors, such as contraction, inner products and traces, all carry over fiberwise to tensor fields. For example, if $\delta_m \in T^1_1(T_mM)$ is the Kronecker delta, then $\delta : M \to T^1_1(M)$; $m \mapsto \delta_m$ is obviously C^∞, and $\delta \in \mathcal{T}^1_1(M)$ is called the ***Kronecker delta***. Similarly, a tensor field of type $(0,s)$ or $(r,0)$ is called ***symmetric***, if it is symmetric at every point.

Metric Tensors. A basic example of a symmetric covariant tensor field is the following.

5.2.12 Definition. *A **weak pseudo-Riemannian metric** on a manifold M is defined to be a tensor field $g \in \mathcal{T}^0_2(M)$ that is symmetric and weakly nondegenerate, that is, such that at each $m \in M$, $g(m)(v_m, w_m) = 0$ for all $w_m \in T_mM$ implies $v_m = 0$. A **strong pseudo-Riemannian metric** is a 2-tensor field that, in addition is **strongly nondegenerate** for all $m \in M$; that is, the map $v_m \mapsto g(m)(v_m, \cdot)$ is an isomorphism of T_mM onto T^*_mM. A weak (resp., strong) pseudo-Riemannian metric is called **weak** (resp., **strong**) **Riemannian** if in addition $g(m)(v_m, v_m) > 0$ for all $v_m \in T_mM$, $v_m \neq 0$.*

It is not hard to show that strong Riemannian manifold is necessarily modeled on a Hilbertizable space; that is, the model space has an equivalent norm arising from an inner product. For finite-dimensional manifolds weak and strong metrics coincide: indeed T_mM and T^*_mM have the same dimension and so a one-to-one map of T_mM to T^*_mM is an isomorphism. It is possible to have weak metrics on a Banach or Hilbert manifold that are not strong. For example, the L^2 inner product on $M = C^0([0,1], \mathbb{R})$ is a weak metric that is not strong. For a similar Hilbert space example, see Exercise 5.2-3.

Any Hilbert space is a Riemannian manifold with a constant metric equal to the inner product. A symmetric bilinear (weakly) nondegenerate two-form on any Banach space provides an example of a (weak) pseudo-Riemannian constant metric. A pseudo-Riemannian manifold used in the theory of special relativity is $\mathbb{R}^4$ with the ***Minkowski pseudo-Riemannian metric***

$$g(x)(v,w) = v^1w^1 + v^2w^2 + v^3w^3 - v^4w^4,$$

where $x, v, w \in \mathbb{R}^4$.

Raising and Lowering Indices. As in the algebraic context of §5.1, pseudo-Riemannian metrics (and for that matter any strongly nondegenerate bilinear tensor) can be used to define associated tensors. Thus the maps $^\sharp, ^\flat$ become vector bundle isomorphisms over the identity $^\flat : TM \to T^*M$, $^\sharp : T^*M \to TM$; $^\sharp$ is the inverse of $^\flat$, where $v_m^\flat = g(m)(v_m, \cdot)$. In particular, they induce isomorphisms of the spaces of sections $^\flat : \mathfrak{X}(M) \to \mathfrak{X}^*(M)$, $^\sharp : \mathfrak{X}^*(M) \to \mathfrak{X}(M)$. In finite dimensions this is the operation of raising and lowering indices. Thus formulas like the ones in Example 5.1.4E should be read pointwise in this context.

Gradients. There is a particular index raising operation that requires special attention.

5.2.13 Definition. *Let M be a pseudo-Riemannian n-manifold with metric g. For $f \in \mathcal{F}(M)$, the vector field defined by* $\operatorname{grad} f = (\mathbf{d}f)^\sharp \in \mathfrak{X}(M)$ *is called the **gradient** of f.*

To find the expression of grad f in local coordinates, we write

$$g_{ij} = g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right), \quad X = X^i \frac{\partial}{\partial x^i}, \quad \text{and} \quad Y = Y^i \frac{\partial}{\partial x^i},$$

so we have

$$\begin{aligned}\left\langle X^\flat, Y\right\rangle &= g(X,Y) = X^i Y^j g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)\\ &= X^i Y^j g_{ij};\end{aligned}$$

that is, $X^\flat = X^i g_{ij} dx^j$. If $\alpha \in \mathfrak{X}^*(M)$ has the coordinate expression $\alpha = \alpha_i dx^i$, we have $\alpha^\sharp = \alpha_i g^{ij} \partial/\partial x^j$ where $[g^{ij}]$ is the inverse of the matrix $[g_{ij}]$. Thus for $\alpha = \mathbf{d}f$, the local expression of the gradient is

$$\operatorname{grad} f = g^{ij} \frac{\partial f}{\partial x^j} \frac{\partial}{\partial x^i}; \quad \text{that is,} \quad (\operatorname{grad} f)^i = g^{ij} \frac{\partial f}{\partial x^j}.$$

If $M = \mathbb{R}^n$ with standard Euclidean metric $g_{ij} = \delta_{ij}$, this formula becomes

$$\operatorname{grad} f = \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^i}; \quad \text{that is,} \quad \operatorname{grad} f = \left(\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n}\right),$$

the familiar expression of the gradient from vector calculus.

Push-forward and Pull-back of Tensor Fields. Now we turn to the effect of mappings and diffeomorphisms on tensor fields.

5.2.14 Definition. *If $\varphi : M \to N$ is a diffeomorphism and $t \in T^r_s(M)$, let $\varphi_* t = (T\varphi)_* \circ t \circ \varphi^{-1}$, be the **push-forward** of t by φ. If $t \in T^r_s(N)$, the **pull-back** of t by φ is given by $\varphi^* t = (\varphi^{-1})_* t$.*

5.2.15 Proposition. *If $\varphi : M \to N$ is a diffeomorphism, and $t \in \mathcal{T}^r_s(M)$, then*

(i) $\varphi_* t \in \mathcal{T}^r_s(N)$;

(ii) $\varphi_* : \mathcal{T}^r_s(M) \to \mathcal{T}^r_s(N)$ *is a linear isomorphism*;

(iii) $(\varphi \circ \psi)_* = \varphi_* \circ \psi_*$; *and*

(iv) $\varphi_*(t \otimes t') = \varphi_* t \otimes \varphi_* t'$, *where $t \in \mathcal{T}^r_s(M)$ and $t' \in \mathcal{T}^r_s(M)$.*

Proof. (i) The differentiability is evident from the composite mapping theorem, together with Proposition 5.2.4. The other three statements are proved fiberwise, where they are consequences of Proposition 5.1.6. ■

As in the algebraic context, the pull-back of covariant tensors is defined even for maps that are not diffeomorphisms. Globalizing Definition 5.1.8 we get the following.

5.2.16 Definition. *If $\varphi : M \to N$ and $t \in \mathcal{T}^0_s(N)$, then $\varphi^* t$, the **pull-back** of t by φ, is defined by*

$$(\varphi^* t)(m)(v_1, \ldots, v_s) = t(\varphi(m))(T_m\varphi(v_1), \ldots, T_m\varphi(v_s))$$

for $m \in M, v_1, \ldots, v_s \in T_m M$.

The next proposition is similar to Proposition 5.2.15 and is proved by globalizing the proof of Proposition 5.1.9.

5.2.17 Proposition. *If $\varphi : M \to N$ is C^∞ and $t \in \mathcal{T}^0_s(N)$, then*

(i) $\varphi^* \in \mathcal{T}^0_s(M)$;

(ii) $\varphi^* : \mathcal{T}^0_s(N) \to \mathcal{T}^0_s(M)$ *is a linear map*;

(iii) $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$ *for* $\psi : N \to P$;

(iv) *if φ is a diffeomorphism them φ^* is an isomorphism with inverse φ^*; and*

(v) $t_1 \in \mathcal{T}^0_{s_1}(N)$, $t_2 \in \mathcal{T}^0_{s_2}(N)$, *then* $\varphi^*(t_1 \otimes t_2) = (\varphi^* t_1) \otimes (\varphi^* t_2)$.

For finite-dimensional manifolds the coordinate expressions of the pull-back and push-forward can be read directly from Propositions 5.1.7 and 5.1.10, taking into account that $T\varphi$ is given locally by the Jacobian matrix. This yields the following.

5.2.18 Proposition. *Let M and N be finite-dimensional manifolds, $\varphi : M \to N$ a C^r map and denote by $y^j = \varphi^j(x^1, \ldots, x^m)$ the local expression of φ relative to charts where $m = \dim(M)$ and $j = 1, \ldots, n = \dim(N)$.*

(i) *If $t \in \mathcal{T}^r_s(M)$ and φ is a diffeomorphism, the coordinates of the push-forward $\varphi_* t$ are*

$$(\varphi_* t)^{i_1 \ldots i_r}_{j_1 \ldots j_s} = \left(\frac{\partial y^{i_1}}{\partial x^{k_1}} \circ \varphi^{-1}\right) \cdots \left(\frac{\partial y^{i_r}}{\partial x^{k_r}} \circ \varphi^{-1}\right) \frac{\partial x^{\ell_1}}{\partial y^{j_1}} \cdots \frac{\partial x^{\ell_s}}{\partial y^{j_s}} t^{k_1 \ldots k_r}_{\ell_1 \ldots \ell_s} \circ \varphi^{-1}.$$

If $t \in \mathcal{T}^0_s(N)$ and φ is a diffeomorphism, the coordinates of the pull-back $\varphi^* t$ are

$$(\varphi^* t)^{i_1 \ldots i_r}_{j_1 \ldots j_s} = \left(\frac{\partial x^{i_1}}{\partial y^{\ell_1}} \circ \varphi\right) \cdots \left(\frac{\partial x^{i_r}}{\partial y^{\ell_r}} \circ \varphi\right) \frac{\partial y^{k_1}}{\partial x^{j_1}} \cdots \frac{\partial y^{k_s}}{\partial x^{j_s}} t^{\ell_1 \ldots \ell_r}_{k_1 \ldots k_s} \circ \varphi.$$

(ii) *If $t \in \mathcal{T}^0_s(N)$ and $\varphi : M \to N$ is arbitrary, the coordinates of the pull-back $\varphi^* t$ are*

$$(\varphi^* t)_{j_1 \ldots j_s} = \frac{\partial y^{k_1}}{\partial x^{j_1}} \cdots \frac{\partial y^{k_s}}{\partial x^{j_s}} t_{k_1 \ldots k_s} \circ \varphi.$$

Notice the similarity between the formulas for coordinate change and pull-back. The situation is similar to the *passive* and *active interpretation* of similarity transformations $\mathbf{PAP}^{-1}$ in linear algebra. Of course it is important not to confuse the two.

5.2.19 Examples.

A. Let $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $\varphi(x, y) = (x + 2y, y)$ and let

$$t = 3x \left(\frac{\partial}{\partial x}\right) \otimes dy + \left(\frac{\partial}{\partial y}\right) \otimes dy \in \mathcal{T}^1_1\left(\mathbb{R}^2\right).$$

The matrix of φ_* on vector fields is

$$\left[\frac{\partial \varphi^i}{\partial x^j}\right] = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

and on forms is

$$\left[\frac{\partial x^i}{\partial \varphi^j}\right] = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}.$$

In other words,

$$\varphi_*\left(\frac{\partial}{\partial x}\right) = \frac{\partial}{\partial x}, \qquad \varphi_*\left(\frac{\partial}{\partial y}\right) = 2\frac{\partial}{\partial x} + \frac{\partial}{\partial y}.$$
$$\varphi_*(dx) = dx - 2\,dy, \qquad \varphi_*(dy) = dy.$$

Noting that $\varphi^{-1}(x, y) = (x - 2y, y)$, we get

$$\begin{aligned}
\varphi_* t &= 3(x - 2y)\varphi_*\left(\frac{\partial}{\partial x}\right) \otimes \varphi_*(dy) + \varphi_*\left(\frac{\partial}{\partial y}\right) \otimes \varphi_*(dy) \\
&= 3(x - 2y)\frac{\partial}{\partial x} \otimes dy + \left(2\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) \otimes dy \\
&= (3x - 6y + 2)\frac{\partial}{\partial x} \otimes dy + \frac{\partial}{\partial y} \otimes dy.
\end{aligned}$$

B. With the same mapping and tensor, we compute $\varphi^* t$. Since

$$\varphi^*\left(\frac{\partial}{\partial x}\right) = \frac{\partial}{\partial x}, \qquad \varphi^*\left(\frac{\partial}{\partial y}\right) = -2\frac{\partial}{\partial x} + \frac{\partial}{\partial y}.$$
$$\varphi^*(dx) = dx + 2\,dy, \qquad \varphi^*(dy) = dy,$$

we have

$$\begin{aligned}
\varphi^* t &= 3(x + 2y)\varphi^*\left(\frac{\partial}{\partial x}\right) \otimes \varphi^*(dy) + \varphi^*\left(\frac{\partial}{\partial y}\right) \otimes \varphi^*(dy) \\
&= 3(x + 3y)\frac{\partial}{\partial x} \otimes dy + \left(-2\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) \otimes dy \\
&= (3x + 6y - 2)\frac{\partial}{\partial x} \otimes dy + \frac{\partial}{\partial y} \otimes dy.
\end{aligned}$$

C. Let $\varphi : \mathbb{R}^3 \to \mathbb{R}^2$, $\varphi(x, y, z) = (2x + z, xyz)$ and $t = (u + 2v)du \otimes du + (u)^2 du \otimes dv \in \mathcal{T}^0_2(\mathbb{R}^2)$. Since

$$\varphi^*(du) = 2dx + dz \quad \text{and} \quad \varphi^*(dv) = yzdx + xzdy + xydz,$$

we have

$$\begin{aligned}
\varphi^* t = {} & (2x + z + 2xyz)(2dx + dz) \otimes (2\,dx + dz) \\
& + (2x + z)^2(2\,dx + dz) \otimes (yz\,dx + xz\,dy + xy\,dz) \\
= {} & 2[4x + 2z + 4xyz + (2x + z)^2 yz]\,dx \otimes dx + 2(2x + z)^2 xz\,dx \otimes dy \\
& + 2[2x + z + 2xyz + (2x + z)^2 xy]\,dx \otimes dz \\
& + [4x + 2z + 4xyz + yz(2x + z)^2]dz \otimes dx \\
& + xz(2x + z)^2\,dz \otimes dy + [2x + z + 2xyz + xy(2x + z)^2]dz \otimes dz.
\end{aligned}$$

D. If $\varphi : M \to N$ represents the deformation of an elastic body and g is a Riemannian metric on N, then $C = \varphi^* g$ is called the ***Cauchy–Green tensor***; in coordinates

$$C_{ij} = \frac{\partial \varphi^\alpha}{\partial x^i} \frac{\partial \varphi^\beta}{\partial x^j} g_{\alpha\beta} \circ \varphi.$$

Thus, C measures how φ deforms lengths and angles. ♦

Alternative Approach to Tensor Fields. Suppose that $\mathcal{F}(M)$ and $\mathfrak{X}(M)$ have been defined. With the "scalar multiplication" $(f, X) \mapsto fX$ defined in Definition 5.2.10, $\mathfrak{X}(M)$ becomes an $\mathcal{F}(M)$-module. That is, $\mathfrak{X}(M)$ is essentially a vector space over $\mathcal{F}(M)$, but the "scalars" $\mathcal{F}(M)$ form only a commutative ring with identity, rather than a field. Define

$$L_{\mathcal{F}(M)}(\mathfrak{X}(M), \mathcal{F}(M)) = \mathcal{X}^*(M)$$

the $\mathcal{F}(M)$-linear mappings on $\mathfrak{X}(M)$, and similarly

$$\mathfrak{T}^r_s(M) = L^{r+s}_{\mathcal{F}(M)}(\mathcal{X}^*(M), \dots, \mathfrak{X}(M); \mathcal{F}(M))$$

the $\mathcal{F}(M)$-multilinear mappings. From Definition 5.2.10, we have a natural mapping $\mathcal{T}^r_s(M) \to \mathfrak{T}^r_s(M)$ which is $\mathcal{F}(M)$-linear.

5.2.20 Proposition. *Let M be a finite-dimensional manifold or be modeled on a Banach space with norm C^∞ away from the origin. Then $\mathcal{T}^r_s(M)$ is isomorphic to $\mathfrak{T}^r_s(M)$ regarded as $\mathcal{F}(M)$-modules and as real vector spaces. In particular, $\mathfrak{X}^*(M)$ is isomorphic to $\mathcal{X}^*(M)$.*

Proof. Consider the map $\mathcal{T}^r_s(M) \to \mathfrak{T}^r_s(M)$ given by

$$\ell(\alpha^1, \dots, \alpha^r, X_1, \dots, X_s)(m) = \ell(m)(\alpha^1(m), \dots, X_s(m)).$$

This map is clearly $\mathcal{F}(M)$-linear. To show it is an isomorphism, given such a multilinear map ℓ, define t by

$$t\,(m)\left(\alpha^1\,(m)\,, \dots, X_s\,(m)\right) = \ell\left(\alpha^1, \dots, X_s\right)(m)\,.$$

To show that t is well-defined we first show that, for each $v_0 \in T_m M$, there is an $X \in \mathfrak{X}(M)$ such that $X(m) = v_0$, and similarly for dual vectors. Let (U, φ) be a chart at m and let $T_m\varphi(v_0) = (\varphi(m), v_0')$. Define $Y \in \mathfrak{X}(U')$ by $Y(u) = (u', v_0')$ on a neighborhood V_1 of $\varphi(m)$, where $w = \varphi(n)$. Extend Y to U' so Y is zero outside V_2, where $\mathrm{cl}(V_1) \subset V_2$, $\mathrm{cl}(V_2) \subset U'$, by means of a bump function. Define X by $X_\varphi = Y$ on U, and $X = 0$ outside U. Then $X(m) = v_0$. The construction is similar for dual vectors.

As in Theorem 4.2.16, $\mathcal{F}(M)$-linearity of ℓ shows that the definition of $t(m)$ is independent of how the vectors v_0 (and corresponding dual vectors) are extended to fields. The tensor field $t(m)$ so defined is C^∞; indeed, using the chart φ, the local representative of t is C^∞ by Supplement 3.4A, since ℓ induces a C^∞ map $M \times T^r_s(M) \to \mathbb{R}$ (by the composite function theorem), which is $(r+s)$-linear at every $m \in M$. If M is finite dimensional this last step of the proof can be simplified as follows. In the chart φ with coordinates $(x^1, \dots, x^n)$,

$$t = t^{i_1 \dots i_r}_{j_1 \dots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial^{i_r}} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s}$$

and all components of t are C^∞ by hypothesis. ■

The preceding proposition can be clearly generalized to the C^k situation. One can also get around the use of a smooth norm on the model space if one assumes that the multilinear maps are ***localizable***, that is, are defined on $\mathfrak{X}^*(U) \times \dots \times \mathfrak{X}(U)$ with values in $\mathcal{F}(U)$ for any open set U in a way compatible with restriction to U. We shall take this point of view in the next section.

The direct sum $\mathcal{T}(M)$ of the $\mathcal{T}^r_s(M)$, including $\mathcal{T}^0_0(M) = \mathcal{F}(M)$, is a real vector space with $\otimes$-product, called the ***tensor algebra*** of M, and if $\varphi : M \to M$ is a diffeomorphism, $\varphi_* : \mathcal{T}(M) \to \mathcal{T}(N)$ is an algebra isomorphism.

The construction of $\mathcal{T}(M)$ and the properties discussed in this section can be generalized to vector bundle valued (r, s)-tensors (resp. tensor fields), that is, elements (resp. sections) of

$$L(T^*M \oplus \cdots \oplus T^*M \oplus TM \oplus \cdots \oplus TM, E),$$

the vector bundle of vector bundle maps from $T^*M \oplus \cdots \oplus TM$ (with r factors of T^*M and s factors of TM) to the vector bundle E, which cover the identity map of the base M.

Exercises

◇ **5.2-1.** Let $\varphi : \mathbb{R}^2 \backslash \{ (0, y) \mid y \in \mathbb{R} \} \to \mathbb{R}^2 \backslash \{ (x, x) \mid x \in \mathbb{R} \}$ be defined by $\varphi(x, y) = (x^3 + y, y)$ and let

$$t = x\frac{\partial}{\partial x} \otimes dx \otimes dy + y\frac{\partial}{\partial y} \otimes dy \otimes dy.$$

Show that φ is a diffeomorphism and compute $\varphi_* t, \varphi^* t$. Endow $\mathbb{R}^2$ with the standard Riemannian metric. Compute the associated tensors of $t, \varphi_* t$, and $\varphi^* t$ as well as their $(1, 1)$ and $(1, 2)$ contractions. What is the trace of the interior product of t with $\partial/\partial x + x\partial/\partial y$?

◇ **5.2-2.** Let $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$, $\varphi(x, y) = (y, x, y + x^2)$ be the deformation of an elastic shell. Compute the Cauchy–Green tensor and its trace.

◇ **5.2-3.** Let H be the set of real sequences $\{a_n\}_{n=1,2,\ldots}$ such that

$$\|a_n\|^2 = \sum_{n \geq 1} n^2 a_n^2 < \infty.$$

Show that H is a Hilbert space. Show that

$$g\,(a, b) = \sum_{n \geq 1} a_n b_n$$

is a weak Riemannian metric on H that is *not* a strong metric.

◇ **5.2-4.** Let (M, g) be a Riemannian manifold and let $N \subset M$ be a submanifold. Define

$$\nu_g(N) = \{ v \in T_n M \mid g(n)(v, u) = 0 \text{ for all} u \in T_n N \text{ and all } n \in N \}.$$

Show that $\nu_g(N)$ is a sub-bundle of $TM|N$ isomorphic to both the normal and conormal bundles $\nu(N)$ and $\mu(N)$ defined in Exercises 3.4-10 and 3.4-11.

5.3 The Lie Derivative: Algebraic Approach

This section extends the Lie derivative $\pounds_X$ from vector fields and functions to the full tensor algebra. We shall do so in two ways. This section does this algebraically and in the next section, it is done in terms of the flow of X. The two approaches will be shown to be equivalent.

Differential Operators. We shall demand certain properties of $\pounds_X$ such as: if t is a tensor field of type (r, s), so is $\pounds_X t$, and $\pounds_X$ should be a derivation for tensor products and contractions. First of all, how should $\pounds_X$ be defined on covector fields? If Y is a vector field and α is a covector field, then the contraction $\alpha \cdot Y$ is a function, so $\pounds_X(\alpha \cdot Y)$ and $\pounds_X Y$ are already defined. (See §4.2.) However, if we require the derivation property for contractions, namely

$$\pounds_X(\alpha \cdot Y) = (\pounds_X \alpha) \cdot Y + \alpha \cdot (\pounds_X Y),$$

then this forces us to define $\pounds_X \alpha$ by

$$(\pounds_X \alpha) \cdot Y = \pounds_X(\alpha \cdot Y) - \alpha \cdot (\pounds_X Y)$$

for all vector fields Y. Since this defines an $\mathcal{F}(M)$-linear map, $\pounds_X \alpha$ is a well-defined covector field. The extension to general tensors now proceeds inductively in the same spirit.

5.3.1 Definition. *A **differential operator** on the full tensor algebra $\mathcal{T}(M)$ of a manifold M is a collection $\{\mathcal{D}^r_s(U)\}$ of maps of $\mathcal{T}^r_s(U)$ into itself for each r and $s \geq 0$ and each open set $U \subset M$, any of which we denote merely $\mathcal{D}$(the r, s and U are to be inferred from the context), such that*

D01. *$\mathcal{D}$ is a **tensor derivation**, or $\mathcal{D}$ **commutes with contractions**, that is, $\mathcal{D}$ is $\mathbb{R}$-linear and if*

$$t \in \mathcal{T}^r_s(M), \quad \alpha_1, \ldots, \alpha_r \in \mathfrak{X}^*(M), \quad \text{and} \quad X_1, \ldots, X_s \in \mathfrak{X}(M),$$

then

$$\begin{aligned} &\mathcal{D}(t(\alpha_1, \ldots, \alpha_r, X_1, \ldots, X_s)) \\ &\qquad = (\mathcal{D}t)(\alpha_1, \ldots, \alpha_r, X_1, \ldots, X_s) \\ &\qquad\quad + \sum_{j=1}^{r} t(\alpha_1, \ldots, \mathcal{D}\alpha_j, \ldots, \alpha_r, X_1, \ldots, X_s) \\ &\qquad\quad + \sum_{k=1}^{s} t(\alpha_1, \ldots, \alpha_r, X_1, \ldots, \mathcal{D}X_k, \ldots, X_s). \end{aligned}$$

D02. *$\mathcal{D}$ is **local**, or is **natural with respect to restrictions**. That is, for $U \subset V \subset M$ open sets, and $t \in \mathcal{T}^r_s(V)$*

$$(\mathcal{D}t)|U = \mathcal{D}(t|U) \in \mathcal{T}^r_s(U)$$

that is, the following diagram commutes

$$\begin{array}{ccc} \mathcal{T}^r_s(V) & \xrightarrow{\ |_U\ } & \mathcal{T}^r_s(U) \\ \mathcal{D}\downarrow & & \downarrow\mathcal{D} \\ \mathcal{T}^r_s(V) & \xrightarrow[\ |_U\ ]{} & \mathcal{T}^r_s(U) \end{array}$$

We do not demand that $\mathcal{D}$ be natural with respect to push-forward by diffeomorphisms. Indeed, several important differential operators, such as the covariant derivative, are not natural with respect to diffeomorphisms, although the Lie derivative is, as we shall see.

5.3.2 Theorem. *Suppose for each open set $U \subset M$ we have maps*

$$\mathcal{E}_U : \mathcal{F}(U) \to \mathcal{F}(U) \quad \text{and} \quad \mathcal{F}_U : \mathfrak{X}(U) \to \mathfrak{X}(U),$$

which are ($\mathbb{R}$-linear) tensor derivations and natural with respect to restrictions. That is

(i) $\mathcal{E}_U(f \otimes g) = (\mathcal{E}_U f) \otimes g + f \otimes \mathcal{E}_U g$ *for* $f, g \in \mathcal{F}(U)$;

(ii) *for* $f \in \mathcal{F}(M)$, $\mathcal{E}_U(f|U) = (\mathcal{E}_M f)|U$;

(iii) $\mathcal{F}_U(f \otimes X) = (\mathcal{E}_U f) \otimes X + f \otimes \mathcal{F}_U X$ *for* $f \in \mathcal{F}(U)$, *and* $X \in \mathfrak{X}(U)$;

(iv) *for* $X \in \mathfrak{X}(M)$, $\mathcal{F}_U f(X|U) = (\mathcal{F}_M X)|U$.

Then there is a unique differential operator $\mathcal{D}$ on $\mathcal{T}(M)$ that coincides with $\mathcal{E}_U$ on $\mathcal{F}(U)$ and with $\mathcal{F}_U$ on $\mathfrak{X}(U)$.

Proof. Since $\mathcal{D}$ must be a tensor derivation, define $\mathcal{D}$ on $\mathfrak{X}^*(U)$ by the formula

$$(\mathcal{D}\alpha) \cdot X = \mathcal{D}(\alpha \cdot X) - \alpha \cdot (\mathcal{D}X) = \mathcal{E}_U(\alpha \cdot X) - \alpha \cdot \mathcal{F}_U X$$

for all $X \in \mathfrak{X}(U)$. By properties (i) and (iii), $\mathcal{D}\alpha$ is $\mathcal{F}(M)$-linear and thus by the remark following Proposition 5.2.20, $\mathcal{D}$ so defined on $\mathfrak{X}^*(U)$ has values in $\mathfrak{X}^*(U)$. Note also that

$$\mathcal{D}(f \otimes \alpha) = (\mathcal{E}f) \otimes \alpha + f \otimes (\mathcal{D}\alpha)$$

for any $\alpha \in \mathfrak{X}^*(U)$, $f \in \mathcal{F}(U)$. This shows that $\mathcal{D}$ exists and is unique on $\mathfrak{X}^*(U)$ (by the Hahn–Banach theorem). Define $\mathcal{D}_U$ on $\mathcal{T}^r_s(U)$ by requiring **D01** to hold:

$$\begin{aligned}(\mathcal{D}_U t)(\alpha_1, \ldots, \alpha_r, X_1, \ldots, X_s) = {} & \mathcal{E}_U(t(\alpha_1, \ldots, \alpha_r, X_1, \ldots, X_s)) \\ & - \sum_{j=1}^{r} t(\alpha_1, \ldots, \mathcal{D}\alpha_j, \ldots, \alpha_r, X_1, \ldots, X_s) \\ & - \sum_{k=1}^{s} t(\alpha_1, \ldots \alpha_r, X_1, \ldots, \mathcal{F}_U X_k, \ldots, X_s).\end{aligned}$$

From (i), (iii), and **D01** for $\mathcal{D}_U$ on $\mathfrak{X}^*(U)$, it follows that $\mathcal{D}_U t$ is an $\mathcal{F}(M)$-multilinear map, that is that $\mathcal{D}_U t \in \mathcal{T}^r_s(U)$ (see the comment following Proposition 5.2.20). The definition of $\mathcal{D}_U$ on $\mathcal{T}^r_s(U)$ uniquely determines $\mathcal{D}_U$ from the property **D01**. Finally, if V is any open subset of U, by (ii) and (iv) it follows that

$$\mathcal{D}_V\,(t|V) = (\mathcal{D}_U t)\,|V.$$

This enables us to define $\mathcal{D}$ on $\mathcal{F}(M)$ by $(\mathcal{D}t)(m) = (\mathcal{D}_U t)(m)$, where U is any open subset of M containing m. Since $\mathcal{D}_U$ is unique, so is $\mathcal{D}$, and so **D02** is satisfied by the construction of $\mathcal{D}$. ■

5.3.3 Corollary. *We have*

(i) $\mathcal{D}(t_1 \otimes t_2) = \mathcal{D}t_1 \otimes t_2 + t_1 \otimes \mathcal{D}t_2$, *and*

(ii) $\mathcal{D}\delta = 0$, *where δ is Kronecker's delta.*

Proof. (i) is a direct application of **D01**. For (ii) let $\alpha \in \mathfrak{X}^*(U)$ and $X \in \mathfrak{X}(U)$ where U is an arbitrary chart domain. Then

$$\begin{aligned}(\mathcal{D}\delta)(\alpha, X) &= \mathcal{D}(\delta(\alpha, X)) - \delta(\mathcal{D}\alpha, X) - \delta(\alpha, \mathcal{D}X) \\ &= \mathcal{D}(\alpha \cdot X) - \mathcal{D}\alpha \cdot X - \alpha \cdot \mathcal{D}X = 0.\end{aligned}$$

Again the Hahn–Banach theorem assures that $\mathcal{D}\delta = 0$ on U, and thus by **D02**, $\mathcal{D}\delta = 0$. ■

The Lie Derivative. Taking $\mathcal{E}_U$ and $\mathcal{F}_U$ to be $\pounds_{X|U}$ we see that the hypotheses of Theorem 5.3.2 are satisfied. Hence we can define a differential operator as follows.

5.3.4 Definition. *If $X \in \mathfrak{X}(M)$, we let $\pounds_X$ be the unique differential operator on $\mathcal{T}(M)$, called the* ***Lie derivative with respect to*** *X, such that $\pounds_X$ coincides with $\pounds_X$ as given on $\mathcal{F}(M)$ and $\mathfrak{X}(M)$ (see Definitions* 4.2.6 *and* 4.2.20*).*

5.3.5 Proposition. *Let $\varphi : M \to N$ be a diffeomorphism and X a vector field on M. Then $\pounds_X$ is* ***natural with respect to push-forward*** *by φ; that is,*

$$\pounds_{\varphi_* X}\varphi_* t = \varphi_* \pounds_X t \quad \textit{for } \mathcal{T}^r_s(M),$$

or the following diagram commutes:

$$\begin{array}{ccc}
\mathcal{T}^r_s(M) & \xrightarrow{\ \varphi_*\ } & \mathcal{T}^r_s(N) \\
\Big\downarrow \pounds_X & & \Big\downarrow \pounds_{\varphi_* X} \\
\mathcal{T}^r_s(M) & \xrightarrow[\ \varphi_*\ ]{} & \mathcal{T}^r_s(N)
\end{array}$$

Proof. For an open set $U \subset M$ define

$$\mathcal{D} : \mathcal{T}^r_s(U) \to \mathcal{T}^r_s(U) \quad \text{by} \quad \mathcal{D}t = \varphi^* \pounds_{\varphi_* X|U}(\varphi_* f),$$

where we use the same symbol φ for $\varphi|U$. By naturality on $\mathcal{F}(U)$ and $\mathfrak{X}(U)$, $\mathcal{D}$ coincides with $\pounds_{X|U}$ on $\mathcal{F}(U)$ and $\mathfrak{X}(U)$. Next, we show that $\mathcal{D}$ is a differential operator. For **D01**, we use the fact that

$$\varphi_*(t(\alpha_1, \ldots, \alpha_r, X_1, \ldots, X_s)) = (\varphi_* t)(\varphi_*\alpha_1, \ldots, \varphi_*\alpha_r, \varphi_* X_1, \ldots, \varphi_* X_s),$$

which follows from the definitions. Then for $X, X_1, \ldots, X_s \in \mathfrak{X}(U)$ and $\alpha_1, \ldots, \alpha_r \in \mathfrak{X}^*(U)$,

$$\begin{aligned}
\mathcal{D}(t(\alpha_1 \ , \ \ldots, \alpha_r, X_1, \ldots, X_s)) &= \varphi^* \pounds_{\varphi_* X}(\varphi_*(t(\alpha_1, \ldots, \alpha_r, X_1, \ldots, X_s))) \\
&= \varphi^* \pounds_{\varphi_* X}((\varphi_* t)(\varphi_*\alpha_1, \ldots, \varphi_*\alpha_r, \varphi_* X_1, \ldots, \varphi_* X_s)) \\
&= \varphi^*[(\pounds_{\varphi_* X}\varphi_* t)(\varphi_*\alpha_1, \ldots, \varphi_*\alpha_r, \varphi_* X_1, \ldots, \varphi_* X_s)) \\
&\quad + \sum_{j=1}^{r} (\varphi_* t)(\varphi_*\alpha_1, \ldots, \pounds_{\varphi_* X}\varphi_*\alpha_j, \ldots, \varphi_*\alpha_r, \varphi_* X_1, \ldots, \varphi_* X_s) \\
&\quad + \sum_{k=1}^{s} (\varphi_* t)(\varphi_*\alpha_1, \ldots, \varphi_*\alpha_r, \varphi_* X_1, \ldots, \pounds_{\varphi_* X}\varphi_* X_k, \ldots, \varphi_* X_s)],
\end{aligned}$$

by **D01** for $\pounds_X$. Since $\varphi^* = (\varphi^{-1})_*$ by Definition 5.2.14, this becomes

$$\begin{aligned}
(\mathcal{D}t)(\alpha_1, \ldots \alpha_r, X_1, \ldots, X_s) &+ \sum_{j=1}^{r} (\alpha_1, \ldots, \mathcal{D}\alpha_j, \ldots, \alpha_r, X_1, \ldots, X_s) \\
&+ \sum_{k=1}^{s} t(\alpha_1, \ldots, \alpha_r, X_1, \ldots, \mathcal{D}X_k, \ldots, X_s).
\end{aligned}$$

For **D02**, let $t \in \mathcal{T}^r_s(M)$ and write

$$\begin{aligned}
\mathcal{D}t|U &= [(\varphi_*)^{-1}\pounds_{\varphi_* X}\varphi_* t]|U = (\varphi_*)^{-1}[\pounds_{\varphi_* X}\varphi_* t]|U \\
&= (\varphi_*)^{-1}\pounds_{\varphi_* X|U}\varphi_* t|U \qquad (\text{by } \mathbf{D02} \text{ for } \pounds_X) \\
&= \mathcal{D}(t|U).
\end{aligned}$$

The result now follows by Theorem 5.3.2. ■

Using the same reasoning, a differential operator that is natural with respect to diffeomorphisms on functions and vector fields is natural on all tensors.

Local Formula for the Lie Derivative. Let us now compute the local formula for $\pounds_X t$ where t is a tensor field of type (r, s). Let $\varphi : U \subset M \to V \subset E$ be a local chart and let X' and t' be the principal parts of the local representatives, $\varphi_* X$ and $\varphi_* t$ respectively. Thus $X' : V \to E$ and $t' : V \to T^r_s(E)$. Recall from §4.2 that the local formulas for the Lie derivatives of functions and vector fields are:

$$(\pounds_X f)'(x) = \mathbf{D}f'(x) \cdot X'(x) \tag{5.3.1}$$

where f' is the local representative of f and

$$(\pounds_X Y)'(x) = \mathbf{D}Y'(x) \cdot X'(x) - \mathbf{D}X'(x) \cdot Y'(x). \tag{5.3.2}$$

In finite dimensions, these become

$$\pounds_X f = X^i \frac{\partial f}{\partial x^i} \tag{5.3.1'}$$

and

$$[X, Y]^i = X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j}. \tag{5.3.2'}$$

Let us first find the local expression for $\pounds_X \alpha$ where α is a one-form. By Proposition 5.3.5, the local representative of $\pounds_X \alpha$ is

$$\varphi_*(\pounds_X \alpha) = \pounds_{\varphi_* X} \varphi_* \alpha,$$

which we write as $\pounds_{X'} \alpha'$ where X' and α' are the principal parts of the local representatives, so $X' : V \to E$ and $\alpha' : V \to E^*$. Let $v \in E$ be fixed and regarded as a constant vector field. Then as $\pounds_X$ is a tensor derivation,

$$\pounds_{X'}(\alpha' \cdot v) = (\pounds_{X'} \alpha') \cdot v + \alpha'(\pounds_{X'} v).$$

By equations (5.3.1) and (5.3.2) this becomes

$$\mathbf{D}(\alpha' \cdot v) \cdot X' = (\pounds_{X'} \alpha') \cdot v - \alpha' \cdot (\mathbf{D}X' \cdot v).$$

Thus,

$$(\pounds_{X'} \alpha') \cdot v = (\mathbf{D}\alpha' \cdot X') \cdot v + \alpha' \cdot (\mathbf{D}X' \cdot v).$$

In the expression $(\mathbf{D}\alpha' \cdot X') \cdot v$, $\mathbf{D}\alpha' \cdot X'$ means the derivative of α' in the direction X'; the resulting element of E^* is then applied to v. Thus we can write

$$\pounds_{X'} \alpha' = \mathbf{D}\alpha' \cdot X' + \alpha' \cdot \mathbf{D}X'. \tag{5.3.3}$$

In finite dimensions, the corresponding coordinate expression is

$$(\pounds_X \alpha)_i v^i = \frac{\partial \alpha^i}{\partial x^j} X^j v^i + \alpha_j \frac{\partial X^j}{\partial x^i} v^i;$$

that is,

$$(\pounds_X \alpha)_i = X^j \frac{\partial \alpha_i}{\partial x^j} + \alpha_j \frac{\partial X^j}{\partial x^i}. \tag{5.3.3'}$$

Now let t be of type (r, s), so $t' : V \to L(E^*, \ldots, E^*, E, \ldots, E; \mathbb{R})$. Let $\alpha^1, \ldots, \alpha^r$ be (constant) elements of E^* and $v_1, \ldots, v_s$ (constant) elements of E. Then again by the derivation property,

$$\begin{aligned} \pounds_{X'}[t'(\alpha^1, \ldots, \alpha^r, v_1, \ldots, v_s)] = (\pounds_{X'} t')(\alpha^1, \ldots, \alpha^r, v_1, \ldots, v_s) \\ + \sum_{i=1}^{r} t'(\alpha^1, \ldots, \pounds_{X'}\alpha^i, \ldots, \alpha^r, v_1, \ldots, v_s) \\ + \sum_{j=1}^{s} t'(\alpha^1, \ldots, \alpha^r, v_1, \ldots, \pounds_{X'} v_j, \ldots, v_s). \end{aligned}$$

Using the local formulas (5.3.1)–(5.3.3) for the Lie derivatives of functions, vector fields, and one-forms, we get

$$\begin{aligned} &(\mathcal{D}t' \cdot X') \cdot (\alpha^1, \ldots, \alpha^r, v_1, \ldots, v_s) \\ &\quad = (\pounds_{X'} t')(\alpha^1, \ldots, \alpha^r, v_1, \ldots, v_s) \\ &\qquad + \sum_{i=1}^{r} t'(\alpha^1, \ldots, \alpha^i \cdot \mathcal{D}X', \ldots, \alpha^r, v_1, \ldots, v_s) \\ &\qquad + \sum_{j=1}^{s} t'(\alpha^1, \ldots, \alpha^r, v_1, \ldots, -\mathcal{D}X' \cdot v_j, \ldots, v_s). \end{aligned}$$

Therefore,

$$\begin{aligned} &(\pounds_{X'} t')(\alpha^1, \ldots, \alpha^r, v_1, \ldots, v_s) \\ &\quad = (\mathcal{D}t' \cdot X')(\alpha^1, \ldots, \alpha^r, v_1, \ldots, v_s) \\ &\qquad - \sum_{i=1}^{r} t'(\alpha^1, \ldots, \alpha^i \cdot \mathcal{D}X', \ldots, \alpha^r, v_1, \ldots, v_s) \\ &\qquad + \sum_{j=1}^{s} t'(\alpha^1, \ldots, \alpha^r, v_1, \ldots, \mathcal{D}X' \cdot v_j, \ldots, v_s). \end{aligned}$$

In components, this reads

$$\begin{aligned} (\pounds_X t)^{i_1 \ldots i_r}_{j_1 \ldots j_s} = X^k \frac{\partial}{\partial x^k} t^{i_1 \ldots i_r}_{j_s \ldots j_s} - \frac{\partial X^{i_1}}{\partial x^\ell} t^{\ell i_2 \ldots i_r}_{j_1 \ldots j_s} - \text{ (all upper indices)} \\ + \frac{\partial X^m}{\partial x^{j_1}} t^{i_1 \ldots i_r}_{m j_2 \ldots j_s} + \text{ (all lower indices)} \end{aligned} \tag{5.3.4}$$

We deduced the component formulas for $\pounds_X t$ in the case of a finite-dimensional manifold as corollaries of the general Banach manifold formulas. Because of their importance, we shall deduce them again in a different manner, without appealing to Proposition 5.3.5. Let

$$t = t^{i_1 \ldots i_r}_{j_1 \ldots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \cdots \otimes dx^{j_s} \in \mathcal{T}^r_s(U),$$

where U is a chart domain on M. If $X = X^k \partial/\partial x^k$, the tensor derivation property can be used to compute $\pounds_X t$. For this we recall that

$$\pounds_X(t^{i_1 \ldots i_r}_{j_1 \ldots j_s}) = X^j \frac{\partial t^{i_1 \ldots i_r}_{j_1 \ldots j_s}}{\partial x^k}$$

and that

$$\pounds_X \frac{\partial}{\partial x^k} = \left[X, \frac{\partial}{\partial x^k}\right] = -\frac{\partial X^i}{\partial x^k}\frac{\partial}{\partial x^i}$$

by the general formula for the bracket components. The formula for $\pounds_X(dx^k)$ is found in the following way. The relation $\delta_i^k = dx^k(\partial/\partial x^i)$ implies by **D01** that

$$\begin{aligned} 0 = \pounds_X\left(dx^k\left(\frac{\partial}{\partial x^i}\right)\right) &= (\pounds_X(dx^k))\left(\frac{\partial}{\partial x^i}\right) + dx^k\left(\left[X, \frac{\partial}{\partial x^i}\right]\right) \\ &= (\pounds_X(dx^k))\left(\frac{\partial}{\partial x^i}\right) + dx^k\left(-\frac{\partial X^\ell}{\partial x^i}\frac{\partial}{\partial x^\ell}\right). \end{aligned}$$

Thus,

$$(\pounds_X(dx^k))\left(\frac{\partial}{\partial x^i}\right) = dx^k\left(\frac{\partial X^\ell}{\partial x^i}\frac{\partial}{\partial x^\ell}\right) = \frac{\partial X^k}{\partial x^i},$$

so

$$\pounds_X(dx^k) = \left(\frac{\partial X^k}{\partial x^i}\right) dx^i.$$

Now one simply applies **D01** and collects terms to get the same local formula for $\pounds_X t$ found in equation (5.3.4). Note especially that

$$\pounds_{\frac{\partial}{\partial x^i}}\left(\frac{\partial}{\partial x^j}\right) = 0 \quad \text{and} \quad \pounds_{\frac{\partial}{\partial x^i}}(dx^j) = 0, \quad \text{for all } i, j.$$

5.3.6 Examples.

A. Compute $\pounds_X t$, where

$$t = x\frac{\partial}{\partial y} \otimes dx \otimes dy + y\frac{\partial}{\partial y} \otimes dy \otimes dy \quad \text{and} \quad X = \frac{\partial}{\partial x} + x\frac{\partial}{\partial y}.$$

Solution

METHOD 1. Note that

(i)
$$\pounds_X t = \pounds_{\partial/\partial x + x\partial/\partial y} t = \pounds_{\partial/\partial x} t + \pounds_{\partial/\partial y} t$$

(ii)
$$\begin{aligned} \pounds_{\partial/\partial x} t &= \pounds_{\partial/\partial x}\left\{x\frac{\partial}{\partial y} \otimes dx \otimes dy + y\frac{\partial}{\partial y} \otimes dy \otimes dy\right\} \\ &= \pounds_{\partial/\partial x}\left(x\frac{\partial}{\partial y} \otimes dx \otimes dy\right) + \pounds_{\partial/\partial x}\left(y\frac{\partial}{\partial y} \otimes dy \otimes dy\right) \\ &= \frac{\partial}{\partial y} \otimes dx \otimes dy + 0. \end{aligned}$$

Now note

$$\pounds_{x\partial/\partial y}\frac{\partial}{\partial y} = 0,$$

$$\pounds_{x\partial/\partial y}\frac{\partial}{\partial x} = -\pounds_{\partial/\partial x}\left(x\frac{\partial}{\partial y}\right) = -\left\{1 \cdot \frac{\partial}{\partial y} + x \cdot 0\right\} = -\frac{\partial}{\partial y},$$

$$\pounds_{x\partial/\partial y} dx = 0, \quad \text{and} \quad \pounds_{x\partial/\partial y} dy = dx.$$

Thus,

(iii)

$$\begin{aligned}
\pounds_{x\partial/\partial y} t &= \pounds_{x\partial/\partial y}\left\{ x\frac{\partial}{\partial y}\otimes dx\otimes dy + y\frac{\partial}{\partial y}\otimes dy\otimes dy\right\}\\
&= \left(0+0+0+x\frac{\partial}{\partial y}\otimes dx\otimes dx\right)\\
&\quad+\left(x\frac{\partial}{\partial y}\otimes dy\otimes dy + 0 + y\frac{\partial}{\partial y}\otimes dx\otimes dy + y\frac{\partial}{\partial y}\otimes dy\otimes dx\right).
\end{aligned}$$

Thus, substituting (ii) and (iii) into (i), we find

$$\begin{aligned}
\pounds_X t &= \frac{\partial}{\partial y}\otimes dx\otimes dy + x\frac{\partial}{\partial y}\otimes dx\otimes dx + x\frac{\partial}{\partial y}\otimes dy\otimes dy\\
&\quad + y\frac{\partial}{\partial y}\otimes dx\otimes dy + y\frac{\partial}{\partial y}\otimes dy\otimes dx\\
&= (y+1)\frac{\partial}{\partial y}\otimes dx\otimes dy + x\frac{\partial}{\partial y}\otimes dx\otimes dx + x\frac{\partial}{\partial y}\otimes dy\otimes dy\\
&\quad + y\frac{\partial}{\partial y}\otimes dy\otimes dx.
\end{aligned}$$

METHOD 2. Using component notation, t is a tensor of type $(1,2)$ whose nonzero components are $t^2_{12}=x$ and $t^2_{22}=y$. The components of X are $X^1=1$ and $X^2=x$. Thus, by the component formula (5.3.4),

$$(\pounds_X t)^i_{jk} = X^k\frac{\partial}{\partial x^k}t^i_{jk} - t^\ell_{jk}\frac{\partial X^i}{\partial x^\ell} + t^i_{mk}\frac{\partial X^m}{\partial x^j} + t^i_{jp}\frac{\partial X^p}{\partial x^k}.$$

The nonzero components are

$$\begin{aligned}
(\pounds_X t)^2_{12} &= 1-0+y+0 = 1+y; &\qquad (\pounds_X t)^2_{22} &= x-0+0+0 = x;\\
(\pounds_X t)^2_{11} &= 0-0+0+x = x; &\qquad (\pounds_X t)^2_{21} &= 0-0+0+y = y,
\end{aligned}$$

and hence

$$\begin{aligned}
\pounds_X t &= (y+1)\frac{\partial}{\partial y}\otimes dx\otimes dy + x\frac{\partial}{\partial y}\otimes dx\otimes dx\\
&\quad + x\frac{\partial}{\partial y}\otimes dy\otimes dy + y\frac{\partial}{\partial y}\otimes dy\otimes dx.
\end{aligned}$$

The two methods thus give the same answer. It is useful to understand both methods since they both occur in the literature, and depending on the circumstances, one may be easier to apply than the other.

B. In Riemannian geometry, vector fields X satisfying $\pounds_X g = 0$ are called ***Killing vector fields***; their geometric significance will become clear in the next section. For now, let us compute the system of equations that the components of a Killing vector field must satisfy. If $X = X^i\partial/\partial x^i$, and $g = g_{ij}dx^i\otimes dx^j$, then

$$\begin{aligned}
\pounds_X g &= (\pounds_X g_{ij})dx^i\otimes dx^j + g_{ij}(\pounds_X dx^i)\otimes dx^j + g_{ij}dx^i\otimes(\pounds_X dx^j)\\
&= X^k\frac{\partial g_{ij}}{\partial x^k}dx^i\otimes dx^j + g_{ij}\frac{\partial X^i}{\partial x^k}dx^k\otimes dx^j + g_{ij}dx^i\otimes\frac{\partial X^j}{\partial x^k}dx^k\\
&= \left\{X^k\frac{\partial g_{ij}}{\partial x^k} + g_{kj}\frac{\partial X^k}{\partial x^i} + g_{ik}\frac{\partial X^k}{\partial x^j}\right\}dx^i\otimes dx^j.
\end{aligned}$$

Note that $\pounds_X g$ is still a symmetric $(0,2)$-tensor, as it must be. Hence X is a Killing vector field iff its components satisfy the following system of n partial differential equations, called ***Killing's equations***

$$X^k \frac{\partial g_{ij}}{\partial x^k} + g_{kj}\frac{\partial X^k}{\partial x^i} + g_{ik}\frac{\partial X^k}{\partial x^j} = 0.$$

C. In the theory of elasticity, if u represents the ***displacement vector field***, the expression $\pounds_u g$ is called the ***strain tensor***. As we shall see in the next section, this is related to the Cauchy–Green tensor $C = \varphi^* g$ by linearization of the deformation φ.

D. Let us show that $\pounds_X$ does not necessarily commute with the formation of associated tensors; for example, that $(\pounds_X t)_{ij} \neq (\pounds_X \tau)_{ij}$, where $t = t^i_j \partial/\partial x^i \otimes dx^j \in T^1_1(M)$ and $\tau = t_{ij} dx^i \otimes dx^j \in T^0_2(M)$ is the associated tensor with components $t_{ij} = g_{ij} t^k_j$. We have from equation (5.3.4)

$$(\pounds_X t)^i_j = X^k \frac{\partial t^i_j}{\partial k^k} - t^k_j \frac{\partial X^i}{\partial x^k} + t^i_k \frac{\partial X^k}{\partial x^j},$$

and so

$$(\pounds_X t)_{ij} g_{i\ell}\left(X^k \frac{\partial t^\ell_j}{\partial x^k} - t^k_j \frac{\partial X^\ell}{\partial x^k} + t^\ell_k \frac{\partial X^k}{\partial x^j}\right).$$

But also from equation (5.3.4)

$$\begin{aligned}
(\pounds_X \tau)_{ij} &= X^k \frac{\partial t_{ij}}{\partial x^k} + t_{\ell_j}\frac{\partial X^\ell}{\partial x^i} + t_{ik}\frac{\partial X^k}{\partial x^j}\\
&= X^k \frac{\partial}{\partial x^k}(g_{i\ell} t^\ell_j) + g_{\ell k} t^k_j \frac{\partial X^\ell}{\partial x^i} + g_{i\ell} t^\ell_k \frac{\partial X^k}{\partial x^j}\\
&= X^k \frac{\partial g_{i\ell}}{\partial x^k} t^\ell_j + X^k g_{i\ell}\frac{\partial t^\ell_j}{\partial x^k} + g_{\ell k} t^k_j \frac{\partial X^\ell}{\partial x^i} + g_{i\ell} t^\ell_k \frac{\partial X^k}{\partial x^j}.
\end{aligned}$$

Thus, to have equality it is necessary and sufficient that

$$X^k \frac{\partial g_{i\ell}}{\partial x^k} t^\ell_j + g_{\ell k} t^k_j \frac{\partial X^\ell}{\partial x^i} + g_{i\ell} t^k_j \frac{\partial X^\ell}{\partial x^k} = 0$$

for all pairs of indices (i,j), which is a nontrivial system of n^2 linear partial differential equations for g_{ij}. If X is a Killing vector field, then

$$g_{\ell k}\frac{\partial X^\ell}{\partial x^i} g_{i\ell}\frac{\partial X^\ell}{\partial x^k} = -X^\ell \frac{\partial g_{ik}}{\partial x^\ell},$$

which substituted in the preceding equation, gives zero. The converse statement is proved along the same lines. In other words, *a necessary and sufficient condition that $\pounds_X$ commute with the formation of associated tensors is that X be a Killing vector field for the pseudo-Riemannian metric g.* ♦

As usual, the development of $\pounds_X$ extends from tensor fields to F-valued tensor fields.

Exercises

◇ **5.3-1.** Let

$$t = xy\frac{\partial}{\partial x} \otimes dx + y\frac{\partial}{\partial y} \otimes dx + \frac{\partial}{\partial x} \otimes dy \in T_1^1\left(\mathbb{R}^2\right).$$

Define the map φ as follows: $\varphi : \{ (x, y) \mid y > 0 \} \to \{ (x, y) \mid x > 0,\ x^2 < y \}$ and $\varphi(x, y) = (ye^x, y^2e^{2x} + y)$. Show that φ is a diffeomorphism and compute trace(t), φ^*t, φ_*t, $\pounds_X t$, $\pounds_X \varphi^* t$, and $L_{\varphi^* X}t$, for $X = y\partial/\partial x + x^2\partial/\partial y$.

◇ **5.3-2.** Verify explicitly that $\pounds_X(t^\flat) \neq (\pounds_X t)^\flat$ where $^\flat$ denotes the associated tensor with both indices lowered, for X and t in Exercise 5.3-1.

◇ **5.3-3.** Compute the coordinate expressions for the Killing equations in $\mathbb{R}^3$ in rectangular, cylindrical, and spherical coordinates. What are the Killing vector fields in $\mathbb{R}^n$?

◇ **5.3-4.** Let (M, g) be a finite dimensional pseudo-Riemannian manifold, and $g^\sharp$ the tensor g with both indices raised. Let $X \in \mathfrak{X}(M)$. Calculate $(\pounds_X g^\sharp)^\flat - \pounds_X g$ in coordinates.

◇ **5.3-5.** If (M, g) is a finite-dimensional pseudo-Riemannian manifold and $f \in \mathcal{F}(M), X \in \mathfrak{X}(M)$, calculate $\pounds_X(\nabla f) - \nabla(\pounds_X f)$.

◇ **5.3-6.** (i) Let $t \in \mathcal{T}_1^1(M)$. Show that there is a unique tensor field $N_t \in \mathcal{T}_2^1(M)$, skew–symmetric in its covariant indices, such that

$$\pounds_{t\cdot X} t - t \cdot \pounds_X t = N_t \cdot X$$

for all $X \in \mathfrak{X}(M)$, where the dots mean contractions, that is $(t \cdot X)^i = t^i_j X^j$, $(t \cdot s)^i_j = t^i_k s^k_j$, where $t, s \in \mathcal{T}_1^1(M)$, and $N_t \cdot X = N^i_{jk} X^k$, where $N_t = N^i_{jk}\partial/\partial x^i \otimes dx^j \otimes dx^k$. N_t is called the ***Nijenhuis tensor***. Generalize to the infinite-dimensional case.

Hint: Show that $N^i_{jk} = t^\ell_k t^i_{j,\ell} - t^\ell_j t^i_{k,\ell} + t^i_\ell t^\ell_{k,j} - t^i_\ell t^\ell_{j,k}$.

(ii) Show that $N_t = 0$ iff

$$[t \cdot X, t \cdot Y] - t \cdot [t \cdot X, Y] = t \cdot [X, t \cdot Y] - t^2 \cdot [X, Y]$$

for all $X, Y \in \mathfrak{X}(M)$, where $t^2 \in \mathcal{T}_1^1(M)$ is the tensor field obtained by the composition $t \circ t$, when t is thought of as a map $t : \mathfrak{X}(M) \to \mathfrak{X}(M)$.

5.4 The Lie Derivative: Dynamic Approach

We now turn to the dynamic interpretation of the Lie derivative. In §4.2 it was shown that $\pounds_X$ acting on an element of $\mathcal{F}(M)$ or $\mathfrak{X}(M)$, respectively, is the time derivative at zero of that element of $\mathcal{F}(M)$ or $\mathfrak{X}(M)$ Lie dragged along by the flow of X. The same situation holds for general tensor fields. Given $t \in \mathcal{T}^r_s(M)$ and $X \in \mathfrak{X}(M)$, we get a curve through $t(m)$ in the fiber over m by using the flow of X. The derivative of this curve is the Lie derivative.

5.4.1 Theorem (Lie Derivative Theorem). *Consider the vector field $X \in \mathfrak{X}^k(M)$, the tensor field $t \in \mathcal{T}^r_s(M)$ both of class C^k, $k \geq 1$ and F_λ be the flow of X. Then on the domain of the flow (see Figure* 5.4.1*), we have*

$$\frac{d}{d\lambda}F_\lambda^* t = F_\lambda^* \pounds_X t.$$

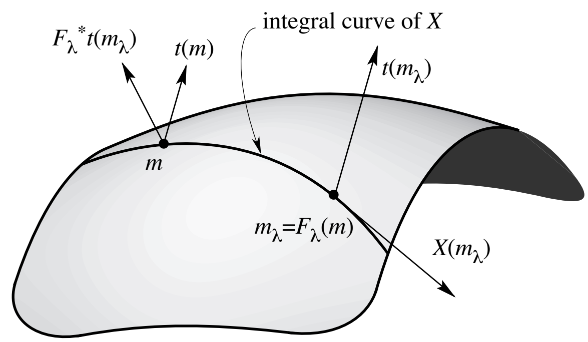

FIGURE 5.4.1. The Lie derivative

Proof. It suffices to show that

$$\left.\frac{d}{d\lambda}\right|_{\lambda=0} F_\lambda^* t = \pounds_X t.$$

Indeed, if this is proved then

$$\frac{d}{d\lambda} F_\lambda^* t = \left.\frac{d}{d\mu}\right|_{\mu=0} F_{\mu+\lambda}^* t = \left.\frac{d}{d\mu}\right|_{\mu=0} F_\lambda^* F_\mu^* t = F_\lambda^* \pounds_X t.$$

Define $\theta_X : \mathcal{T}(M) \to \mathcal{T}(M)$ by

$$\theta_X(t)(m) = \left.\frac{d}{d\lambda}\right|_{\lambda=0} (F_\lambda^* t)(m).$$

Note that $\theta_X(t)$ is a smooth tensor field of the same type as t, by smoothness of t and F_λ. (We suppress the notational clutter of restricting to the domain of the flow.) Let us apply Theorem 5.3.2. Clearly θ_X is $\mathbb{R}$-linear and is natural with respect to restrictions. It is a tensor derivation from the product rule for derivatives and the relation

$$(\varphi^* t)(\varphi^* \alpha^1, \ldots, \varphi^* \alpha^r, \varphi^* X_1, \ldots, \varphi^* X_s) = \varphi^*(t(\alpha^1, \ldots, \alpha^r, X_1, \ldots, X_s))$$

for φ a diffeomorphism. Hence θ_X is a differential operator. It remains to show that θ_X coincides with $\pounds_X$ on $\mathcal{F}(M)$ and $\mathfrak{X}(M)$. For $f \in \mathcal{F}(M)$, and $X \in \mathfrak{X}(M)$, we have

$$\theta_X f = \left.\frac{d}{d\lambda}\right|_{\lambda=0} F_\lambda^* f = \pounds_X f \quad \text{and} \quad \theta_X Y = \left.\frac{d}{d\lambda}\right|_{\lambda=0} F_\lambda^* Y = [X, Y]$$

by Theorems 4.2.10 and 4.2.19, respectively. By Theorem 5.3.2 and Definition 5.3.4, $\theta_X t = \pounds_X t$ for all $t \in \mathcal{T}(M)$. ■

This theorem can also be verified in finite dimensions by a straightforward coordinate computation. See Exercise 5.4-1.

The identity in this theorem relating flows and Lie derivatives is so basic, some authors like to take it as the *definition* of the Lie derivative (see Exercise 5.4-3).

5.4.2 Corollary. *If $t \in \mathcal{T}(M)$, $\pounds_X t = 0$ iff t is constant along the flow of X. That is, $t = F_\lambda^* t$.*

As an application of Theorem 5.4.1, let us generalize the naturality of $\pounds_X$ with respect to diffeomorphisms. As remarked in §5.2, the pull-back of covariant tensor fields makes sense even when the mapping is not a diffeomorphism. It is thus natural to ask whether there is some analogue of Proposition 5.3.5 for pull-backs with no invertibility assumption on the mapping φ. Of course, the best one can hope for, since vector fields can be operated upon only by diffeomorphisms, is to replace the pair X, $\varphi_* X$ be a pair of φ-related vector fields.

5.4.3 Proposition. *Let $\varphi : M \to N$ be C^∞, $X \in \mathfrak{X}(M)$, $Y \in \mathfrak{X}(N)$, $X \sim_\varphi Y$, and $t \in \mathcal{T}^0_s(N)$. Then $\varphi^*(\pounds_Y t) = \pounds_X \varphi^* t$.*

Proof. Recall from Proposition 4.2.4 that $X \sim_\varphi Y$ iff $G_\lambda \circ \varphi = \varphi \circ F_\lambda$, where F_λ and G_λ are the flows of X and Y, respectively. Thus by Theorem 5.4.1,

$$\begin{aligned} \pounds_X(\varphi^* t) &= \left.\frac{d}{d\lambda}\right|_{\lambda=0} F_\lambda^* \varphi^* t = \left.\frac{d}{d\lambda}\right|_{\lambda=0} (\varphi \circ F_\lambda)^* t \\ &= \left.\frac{d}{d\lambda}\right|_{\lambda=0} (G_\lambda \circ \varphi)^* t = \left.\frac{d}{d\lambda}\right|_{\lambda=0} \varphi^* G_\lambda^* t \\ &= \varphi^* \left.\frac{d}{d\lambda}\right|_{\lambda=0} G_\lambda^* t = \varphi^*(\pounds_Y t). \end{aligned}$$

■

As with functions and vector fields, the Lie derivative can be generalized to include time-dependent vector fields.

5.4.4 Theorem (First Time-dependent Lie Derivative Theorem)**.** *Let $X_\lambda \in \mathfrak{X}^k(M)$, $k \geq 1$, for $\lambda \in \mathbb{R}$ and suppose that $X(\lambda, m)$ is continuous in (λ, m). Letting $F_{\lambda,\mu}$ be the evolution operator for X_λ, we have*

$$\frac{d}{d\lambda} F_{\lambda,\mu}^* t = F_{\lambda,\mu}^*(\pounds_{X_\lambda} t)$$

where $t \in \mathcal{T}^r_s(M)$ is of class C^k.

Warning. It is *not* generally true for time-dependent vector fields that the right hand-side in the preceding display equals

$$\pounds_{X_\lambda} F_{\lambda,\mu}^* t.$$

Proof. As in Theorem 5.4.1, it is enough to prove the formula at $\lambda = \mu$ where $F_{\lambda,\lambda} =$ identity, for then

$$\begin{aligned} \frac{d}{d\lambda} F_{\lambda,\mu}^* t &= \left.\frac{d}{d\rho}\right|_{\rho=\lambda} (F_{\rho,\lambda} \circ F_{\lambda,\mu})^* t = F_{\lambda,\mu}^* \left.\frac{d}{d\rho}\right|_{\rho=\lambda} F_{\rho,\lambda}^* t \\ &= F_{\lambda,\mu}^* \pounds_{X_\lambda} t. \end{aligned}$$

As in Theorem 5.4.1,

$$\theta_{X_\lambda} t = \left.\frac{d}{d\lambda}\right|_{\lambda=\mu} F_{\lambda,\mu}^* t$$

is a differential operator that coincides with $\pounds_{X_\lambda}$ on $\mathcal{F}(M)$ and on $\mathfrak{X}(M)$ by Theorem 4.2.31. Thus by Theorem 5.3.2, $\theta_{X_\lambda} = \pounds_{X_\lambda}$ on all tensors. ■

Let us generalize the relationship between Lie derivatives and flows one more step. Call a smooth map $t : \mathbb{R} \times M \to \mathcal{T}^r_s(M)$ satisfying $t_\lambda(m) = t(\lambda, m) \in (T_m M)^r_s$ a ***time-dependent tensor field***. Theorem 5.4.4 generalizes to this context as follows.

5.4.5 Theorem (Second Time-dependent Lie Derivative Theorem). *Let t_λ be a C^k time-dependent tensor, and X_λ be as in Theorem* 5.4.4, $k \geq 1$, *and denote by $F_{\lambda,\mu}$ the evolution operator of X_λ. Then*

$$\frac{d}{d\lambda} F_{\lambda,\mu}^* t_\lambda = F_{\lambda,\mu}^* \left(\frac{\partial t_\lambda}{\partial \lambda} + \pounds_{X_\lambda} t_\lambda \right)$$

Proof. By the product rule for derivatives and Theorem 5.4.4 we get

$$\begin{aligned} \left.\frac{d}{d\lambda}\right|_{\lambda=\sigma} F_{\lambda,\mu}^* t_\lambda &= \left.\frac{d}{d\lambda}\right|_{\lambda=\sigma} F_{\lambda,\mu}^* t_\sigma + F_{\sigma,\mu}^* \left.\frac{dt_\lambda}{d\lambda}\right|_{\lambda=\sigma} \\ &= F_{\sigma,\mu}^* (\pounds_{X_\sigma} t_\sigma) + F_{\sigma,\mu}^* \left.\frac{dt_\lambda}{d\lambda}\right|_{\lambda=\sigma}. \end{aligned}$$

■

5.4.6 Examples.

A. If g is a pseudo-Riemannian metric on M, the Killing equations are $\pounds_X g = 0$ (see Example 5.3.6B). By Corollary 5.4.2 this says that $F_\lambda^* g = g$, where F_λ is the flow of X, that is that the flow of X consists of isometries.

B. In elasticity, the vanishing of the strain tensor means, by Example A, that the body moves as a rigid body. ♦

We close this section with an important technique based on the dynamic approach to the Lie derivative, called the ***Lie transform method***. It has been used already in the proof of the Frobenius theorem (§4.4) and we shall see it again in Chapters 6 and 9. The method is also used in the theory of *normal forms* (cf. Takens [1974], Guckenheimer and Holmes [1983], and Golubitsky and Schaeffer [1985]).

5.4.7 Example (The Lie Transform Method). Let two tensor fields t_0 and t_1 be given on a smooth manifold M. We say they are ***locally equivalent*** at $m_0 \in M$ if there is a diffeomorphism φ of a neighborhood of m_0 to itself, such that $\varphi^* t_1 = t_0$. *One way to show that t_0 and t_1 are equivalent is to join them with a curve $t(\lambda)$ satisfying $t(0) = t_0$, $t(1) = t_1$ and to seek a curve of local diffeomorphisms φ_λ such that $\varphi_0 =$* identity *and*

$$\varphi_\lambda^* t(\lambda) = t_0, \quad \lambda \in [0,1].$$

If this is done, $\varphi = \varphi_1$ is the desired diffeomorphism. A way to find the curve of diffeomorphisms φ_λ satisfying the relation above is to *solve the equation*

$$\pounds_{X_\lambda} t(\lambda) + \frac{d}{d\lambda} t(\lambda) = 0$$

for X_λ. If this is possible, let $\varphi_\lambda = F_{\lambda,0}$, where $F_{\lambda,\mu}$ is the evolution operator of the time-dependent vector field X_λ. Then by Theorem 5.4.5 we have

$$\frac{d}{d\lambda} \varphi_\lambda^* t(\lambda) = \varphi_\lambda^* \left(\pounds_{X_\lambda} t(\lambda) + \frac{d}{d\lambda} t(\lambda) \right) = 0$$

so that $\varphi_\lambda^* t(\lambda) = \varphi_0^*(0) = t_0$. If we choose X_λ so $X_\lambda(m_0) = 0$, then φ_λ exists for a time ≥ 1 by Corollary 4.1.25 and $\varphi_\lambda(m_0) = m_0$.

One often takes $t(\lambda) = (1-\lambda)t_0 + \lambda t_1$. Also, in applications this method is not always used in exactly this way since the algebraic equation for X_λ might be hard to solve. We shall see this happen in the proof of the Poincaré lemma 6.4.14. The reader should now also look back at the Frobenius theorem 4.4.7 and recognize the spirit of the Lie transform method in its proof. ♦

We shall next prove a version of the classical Morse lemma in infinite dimensions using the method of Lie transforms. The proof below is due to Golubitsky and Marsden [1983]; see Palais [1969] and Tromba [1976] for the original proofs; Palais' proof is similar in spirit to the one we give.

5.4.8 Lemma (The Morse–Palais–Tromba Lemma). *Let* $\mathbf{E}$ *be a Banach space and* $\langle\,,\,\rangle$ *a weakly nondegenerate, continuous, symmetric bilinear form on* $\mathbf{E}$. *Let* $h : U \to \mathbb{R}$ *be* C^k, $k \geq 3$, *where* U *is open in* $\mathbf{E}$, *and let* $u_0 \in U$ *satisfy* $h(u_0) = 0$, $\mathbf{D}h(u_0) = 0$. *Let* $B = \mathbf{D}^2 h(u_0) : \mathbf{E} \times \mathbf{E} \to \mathbb{R}$. *Assume that there is a linear isomorphism* $T : \mathbf{E} \to \mathbf{E}$ *such that* $B(u, v) = \langle Tu, v\rangle$ *for all* $u, v \in \mathbf{E}$ *and that* h *has a* C^{k-1} *gradient*

$$\langle \nabla h(y), u\rangle = \mathbf{D}h(y) \cdot u.$$

Then there is a local C^{k-2} *diffeomorphism* φ *of* $\mathbf{E}$ *with* $\varphi(u_0) = u_0$, $\mathbf{D}\varphi(u_0) = I$, *and*

$$h(\varphi(x)) = \frac{1}{2} B(x - u_0, x - u_0).$$

Proof. Symmetry of B implies that T is self-adjoint relative to $\langle\,,\,\rangle$. Let

$$f(y) = \left(\frac{1}{2}\right) B\left(y - u_0, y - u_0\right),$$

$h_1 = h$, and $h_\lambda = f + \lambda p$, where

$$p(y) = h(y) - \left(\frac{1}{2}\right) B\left(y - u_0, y - u_0\right)$$

is C^k and satisfies $p(u_0) = 0$, $\mathbf{D}p(u_0) = 0$, and $\mathbf{D}^2 p(u_0) = 0$. We apply the Lie transform method to h_λ. Thus we have to solve the following equation for a C^{k-2} vector field X_λ

$$\pounds_{X_\lambda} h_\lambda + \frac{dh_\lambda}{d\lambda} = 0, \quad X_\lambda(u_0) = 0. \tag{5.4.1}$$

Then $\varphi_1^* h = f$, where $\varphi_\lambda = F_{\lambda,0}$, for $F_{\lambda,\mu}$ the evolution operator of X_λ, and hence $\varphi = \varphi_1$ is a C^{k-2} diffeomorphism of a neighborhood of u_0 satisfying $\varphi_1(u_0) = u_0$. If we can prove that $\mathbf{D}\varphi_1(u_0) = I$, $\varphi_1 = \varphi$ will be the desired diffeomorphism.

To solve equation (5.4.1), differentiate $\mathbf{D}p(x) \cdot e = \langle \nabla p(x), e\rangle$ with respect to x and use the symmetry of the second derivative to conclude that $\mathbf{D}\nabla p(x)$ is symmetric relative to $\langle\,,\,\rangle$. Therefore,

$$\begin{aligned} \mathbf{D}p(x) \cdot e = \langle \nabla p(x), e\rangle &= \left\langle \int_0^1 \mathbf{D}\nabla p(u_0 + \tau(x - u_0)) \cdot (x - u_0) d\tau, e \right\rangle \\ &= \langle T(x - u_0), R(x) \cdot e\rangle \end{aligned} \tag{5.4.2}$$

where $R : U \to L(\mathbf{E}, \mathbf{E})$ is the C^{k-2} map given by

$$R(x) = T^{-1} \int_0^1 \mathbf{D}p(u_0 + \tau(x - u_0)) \cdot (x - u_0)\, d\tau$$

which satisfies $R(u_0) = 0$. Thus $p(y)$ has the expression

$$p(y) = \int_0^1 \mathbf{D}p(u_0 + \tau(y - u_0)) \cdot (y - u_0) d\tau = -\langle T(y - u_0), X(y)\rangle$$

where $X : U \to \mathbf{E}$ is the C^{k-2} vector field given by

$$X(y) = -\int_0^1 \tau R(u_0 + \tau(y - u_0)) \cdot (y - u_0)\, d\tau$$

which satisfies $X(u_0) = 0$ and $\mathbf{D}X(u_0) = I$. Therefore

$$(\pounds_{X_\lambda} h_\lambda)(y) = \mathbf{D}h_\lambda(y) \cdot X_\lambda(y) = B(y - u_0, X_\lambda(y)) + \lambda \mathbf{D}p(y) \cdot X_\lambda(y)$$
$$= \langle T(y - u_0), (I + \lambda R(y)) \cdot X_\lambda(y) \rangle \qquad \text{by (equation 5.4.2)}$$

so that the equation (5.4.1) becomes,

$$\langle T(y - u_0), (I + \lambda R(y)) \cdot X_\lambda(y) \rangle = \langle T(y - u_0), X(y) \rangle \, .$$

Since $R(u_0) = 0$, there exists a neighborhood of u_0, such that the norm of $\lambda R(y)$ is < 1 for all $\lambda \in [0, 1]$. Thus for y in this neighborhood, $I + \lambda R(y)$ can be inverted and we can take $X_\lambda(y) = (I + \lambda R(y))^{-1} X(y)$ which is a C^{k-2} vector field defined for all $\lambda \in [0, 1]$ and which satisfies $X_\lambda(u_0) = 0$, $\mathbf{D}X_\lambda(u_0) = 0$. Differentiating the relation $(d/d\lambda)\varphi_\lambda(u) = X_\lambda(\varphi_\lambda(u))$ in u at u_0 and using $\varphi_\lambda(u_0) = u_0$ yields

$$\frac{d}{d\lambda} \mathbf{D}\varphi_\lambda\left(u_0\right) = \mathbf{D}X_\lambda\left(\varphi_\lambda\left(u_0\right)\right) \circ \mathbf{D}\varphi_\lambda\left(u_0\right) = \mathbf{D}X_\lambda\left(u_0\right) \circ \mathbf{D}\varphi_\lambda\left(u_0\right) = 0,$$

that is $\mathbf{D}\varphi_\lambda(u_0)$ is constant in $\lambda \in [0, 1]$. Since it equals I at $\lambda = 0$, it follows that $\mathbf{D}\varphi_\lambda(u_0) = I$. ■

5.4.9 Lemma (The Classical Morse Lemma). *Let $h : U \to \mathbb{R}$ be C^k, $k \geq 3$, U open in $\mathbb{R}^n$, and let $u \in U$ be a nondegenerate critical point of h, that is $h(u) = 0$, $\mathbf{D}h(u) = 0$ and the symmetric bilinear form $\mathbf{D}^2 h(u)$ on $\mathbb{R}^n$ is nondegenerate. Then there is a local C^{k-2} diffeomorphism ψ of $\mathbb{R}^n$ fixing u such that*

$$h\left(\psi(x)\right) = \frac{1}{2}\left[\left(x^1 - u^1\right)^2 + \cdots + \left(0x - u^{n-i}\right)^2 - \left(x^{n-i+1} - u^{n-i+1}\right)^2 - \cdots - \left(x^n - u^n\right)^2\right].$$

Proof. In Lemma 5.4.8, take $\langle \, , \rangle$ to be the dot-product in $\mathbb{R}^n$ to find a local C^{k-2} diffeomorphism on $\mathbb{R}^n$ fixing u_0 such that

$$h(\varphi(x)) = (1/2)\mathbf{D}^2 h(u)(x - u, x - u).$$

Next, apply the Gram–Schmidt procedure to find a basis of $\mathbb{R}^n$ in which the matrix of $\mathbf{D}^2 h(u)$ is diagonal with entries ± 1 (see Proposition 6.2.9 for a review of the proof of the existence of such a basis). If i is the number of -1's (the ***index***), let φ be the composition of ψ with the linear isomorphism determined by the change of an arbitrary basis of $\mathbb{R}^n$ to the one above. ■

Exercises

◇ **5.4-1.** Verify Theorem 5.4.1 by a coordinate computation as follows. Let $F_\lambda(x) = (y^1(\lambda, x), \ldots, y^n(\lambda, x))$ so that $\partial y^i / \partial \lambda = X^i(y)$ and $\partial y^i / \partial x^j$ satisfy the variational equation

$$\frac{\partial}{\partial \lambda} \frac{\partial y^i}{\partial x^j} = \frac{\partial X^i}{\partial x^k} \frac{\partial y^k}{\partial x^j}.$$

Then write

$$(F_\lambda^* t)^{a_1 \ldots a_r}_{b_1 \ldots b_s} = \frac{\partial x^{a_1}}{\partial y^{i_1}} \cdots \frac{\partial x^{a_r}}{\partial y^{i_r}} \frac{\partial y^{j_1}}{\partial x^{b_1}} \cdots \frac{\partial y^{j_s}}{\partial x^{b_s}} t^{i_1 \ldots i_r}_{j_1 \ldots j_s}.$$

Differentiate this in λ at $\lambda = 0$ and obtain the coordinate expression (5.3.4) of Section 5.3 for $\pounds_X t$.

◇ **5.4-2.** Carry out the proof outlined in Exercise 5.4-1 for time-dependent vector fields.

◇ **5.4-3.** Starting with Theorem 5.4.1 as the definition of $\pounds_X t$, check that $\pounds_X$ satisfies **DO1**, **DO2** and the properties (i)–(iv) of Theorem 5.3.2.

◇ **5.4-4.** Let $\mathcal{C}$ be a contraction operator mapping $\mathcal{T}^r_s(M)$ to $\mathcal{T}^{r-1}_{s-1}(M)$. Use both Theorem 5.4.1 and **DO1** to show that $\pounds_X(\mathcal{C}t) = \mathcal{C}(\pounds_X t)$.

◇ **5.4-5.** Extend Theorem 5.4.1 to **F**-valued tensors.

◇ **5.4-6.** Let $f(y) = (1/2)y^2 - y^3 + y^5$. Use the Lie transform method to show that there is a local diffeomorphism φ, defined in a neighborhood of $0 \in \mathbb{R}$ such that $(f \circ \varphi)(x) = x^2/2$.

◇ **5.4-7.** Let $\mathbf{E} = \ell^2(\mathbb{R})$, let

$$\langle x, y\rangle = \sum_{n=1}^{\infty} \frac{1}{n} x_n y_n \quad \text{and} \quad h(x) = \frac{1}{2}\sum_{n=1}^{\infty} \frac{1}{n} x_n^2 = \frac{1}{3}\sum_{n=1}^{\infty} x_n^3.$$

Show that h vanishes on $(0, 0, \ldots, 3/2n, 0, \ldots)$ which $\to 0$ as $n \to \infty$, so the conclusion of the Morse lemma fails. What hypothesis in Lemma 5.4.8 fails?

◇ **5.4-8** (Buchner, Marsden, and Schecter [1983b]). In the notation of Exercise 2.4-15, show that f has a sequence of critical points approaching 0, so the Morse lemma fails. (The only missing hypothesis is that ∇h is C^1.)

5.5 Partitions of Unity

A partition of unity is a technical device that is often used to piece smooth local tensor fields together to form a smooth global tensor field. Partitions of unity will be useful for studying integration; in this section they are used to study when a manifold admits a Riemannian metric.

5.5.1 Definition. *If t is a tensor field on a manifold M, the* ***carrier*** *of t is the set of $m \in M$ for which $t(m) \neq 0$, and is denoted* $\operatorname{carr}(t)$. *The* ***support*** *of t, denoted* $\operatorname{supp}(t)$, *is the closure of* $\operatorname{carr}(t)$. *We say t has* ***compact support*** *if* $\operatorname{supp}(t)$ *is compact in M. An open set $U \subset M$ is called a C^r* ***carrier*** *if there exists an $f \in \mathcal{F}^r(M)$, such that $f \geq 0$ and $U = \operatorname{carr}(f)$. A collection of subsets $\{C_\alpha\}$ of a manifold M (or, more generally, a topological space) is called* ***locally finite*** *if for each $m \in M$, there is a neighborhood U of m such that $U \cap C_\alpha = \varnothing$ except for finitely many indices α.*

5.5.2 Definition. *A* ***partition of unity*** *on a manifold M is a collection $\{(U_i, g_i)\}$, where*

(i) *$\{U_i\}$ is a locally finite open covering of M;*

(ii) *$g_i \in \mathcal{F}(M)$, $g_i(m) \geq 0$ for all $m \in M$, and* $\operatorname{supp}(g_i) \subset U_i$ *for all i;*

(iii) *for each $m \in M$, $\Sigma_i g_i(m) = 1$. (By* (i)*, this is a finite sum.)*

If $\mathcal{A} = \{(V_\alpha, \varphi_\alpha)\}$ is an atlas on M, a ***partition of unity subordinate*** *to $\mathcal{A}$ is a partition of unity $\{(U_i, g_i)\}$ such that each open set U_i is a subset of a chart domain $V_{\alpha(i)}$. If any atlas $\mathcal{A}$ has a subordinate partition of unity, we say M* ***admits partitions of unity****.*

Occasionally one works with C^k partitions of unity. They are defined in the same way except g_i are only required to be C^k rather than C^∞.

5.5.3 Theorem (Patching Construction). *Let M be a manifold with an atlas $\mathcal{A} = \{(V_\alpha, \varphi_\alpha)\}$ where $\varphi_\alpha : V_\alpha \to V'_\alpha \subset \mathbf{E}$ is a chart. Let t_α be a C^k tensor field, $k \geq 1$, of fixed type (r, s) defined on V'_α for each α, and assume that there exists a partition of unity $\{(U_i, g_i)\}$ subordinate to $\mathcal{A}$. Let t be defined by*

$$t(m) = \sum_i g_i \varphi^*_{\alpha(i)} t_{\alpha(i)}(m),$$

a finite sum at each $m \in M$. Then t is a C^k tensor field of type (r, s) on M.

Proof. Since $\{U_i\}$ is locally finite, the sum at every point is a finite sum, and thus $t(m)$ is a type (r,s) tensor for every $m \in M$. Also, t is C^k since the local representative of t in the chart $(V_{\alpha(i)}, \varphi_{\alpha(i)})$ is $\Sigma_j (g_i \circ \varphi_{\alpha(j)}^{-1}) t_{\alpha(j)}$, the summation taken over all indices j such that $V_{\alpha(i)} \cap V_{\alpha(j)} \neq \varnothing$; by local finiteness the number of these j is finite. ■

Clearly this construction is not unique; it depends on the choices of the indices $\alpha(i)$ such that $U_i \subset V_{\alpha(i)}$ and on the functions g_i. As we shall see later, under suitable hypotheses, one can always construct partitions of unity; again the construction is not unique. The same construction (and proof) can be used to patch together local sections of a vector bundle into a global section when the base is a manifold admitting partitions of unity subordinate to any open covering.

To discuss the existence of partitions of unity and consequences thereof, we need some topological preliminaries.

5.5.4 Definition. *Let S be a topological space. A covering $\{U_\alpha\}$ of S is called a **refinement** of a covering $\{V_i\}$ if for every U_α there is a V_i such that $U_\alpha \subset V_i$. A topological space is called **paracompact** if every open covering of S has a locally finite refinement of open sets, and S is Hausdorff.*

5.5.5 Proposition. *Second-countable, locally compact Hausdorff spaces are paracompact.*

Proof. By second countability and local compactness of S, there exists a sequence $O_1, \ldots, O_n, \ldots$ of open sets with $\mathrm{cl}(O_n)$ compact and $\bigcup_{n\in\mathbb{N}} O_n = S$. Let $V_n = O_1 \cup \cdots \cup O_n$, $n = 1, 2, \ldots$ and put $U_1 = V_1$. Since $\{V_n\}$ is an open covering of S and $\mathrm{cl}(U_1)$ is compact,

$$\mathrm{cl}(U_1) \subset V_{i_1} \cup \cdots \cup V_{i_r}.$$

Put

$$U_2 = V_{i_1} \cup \cdots \cup V_{i_r};$$

then $\mathrm{cl}(U_2)$ is compact. Proceed inductively to show that S is the countable union of open sets U_n such that $\mathrm{cl}(U_n)$ is compact and $\mathrm{cl}(U_n) \subset U_{n+1}$. If W_α is a covering of S by open sets, and $K_n = \mathrm{cl}(U_n) \backslash U_{n-1}$, then we can cover K_n by a finite number of open sets, each of which is contained in some $W_\alpha \cap U_{n+1}$, and is disjoint from $\mathrm{cl}(U_{n-2})$. The union of such collections yields the desired refinement of $\{W_\alpha\}$. ■

Another class of paracompact spaces are the metrizable spaces (see Lemma 5.5.15 in Supplement 5.5A). In particular, Banach spaces are paracompact.

5.5.6 Proposition. *Every paracompact space is normal.*

Proof. We first show that if A is closed and $u \in S\backslash A$, there are disjoint neighborhoods of u and A (regularity). For each $v \in A$, let U_u, V_v be disjoint neighborhoods of u and v. Let W_α be a locally finite refinement of the covering $\{\, V_v, S\backslash A \mid v \in A \,\}$, and $V = \bigcup W_\alpha$, the union over those α with $W_\alpha \cap A \neq \varnothing$. A neighborhood U_0 of u meets a finite number of W_α. Let U denote the intersection of U_0 and the corresponding U_u. Then V and U are the required neighborhoods. The case for two closed sets proceeds somewhat similarly, so we leave the details for the reader. ■

Later we shall give general theorems on the existence of partitions of unity. However, there is a simple case that is commonly used, so we present it first.

5.5.7 Theorem. *Let M be a second-countable (Hausdorff) n-manifold. Then M admits partitions of unity.*

Proof. The proof of Proposition 5.5.5 shows the following. Let M be an n-manifold and $\{W_\alpha\}$ be an open covering. Then there is a locally finite refinement consisting of charts (V_i, φ_i) such that $\varphi_i(V_i)$ is the disk of radius 3, and such that $\varphi_i^{-1}(D_1(0))$ cover M, where $D_1(0)$ is the unit disk, centered at the origin in the model space. Now let $\mathcal{A}$ be an atlas on M and let $\{(V_i, \varphi_i)\}$ be a locally finite refinement with these properties. From Lemma 4.2.13, there is a nonzero function $h_i \in \mathcal{F}(M)$ whose support lies in V_i and $h_j \geq 0$. Let

$$g_i(u) = \frac{h_i(u)}{\Sigma_i h_i(u)}$$

(the sum is finite). These are the required functions. ■

If $\{V_\alpha\}$ is an open covering of M, we can always find an atlas $\mathcal{A} = \{(U_i, \varphi_i)\}$ such that $\{U_i\}$ is a refinement of $\{V_\alpha\}$ since the atlases generate the topology. Thus, if M admits partitions of unity, we can find partitions of unity subordinate to any open covering.

The case of C^0-partitions of unity differs drastically from the smooth case. Since we are primarily interested in this latter case, we summarize the topological situation, without giving the proofs.

1. *If S is a Hausdorff space, the following are equivalent*:

(i) *S is normal.*

(ii) (Urysohn's lemma.) *For any two closed nonempty disjoint sets A, B there is a continuous function $f : S \to [0,1]$ such that $f(A) = 0$ and $f(B) = 1$.*

(iii) (Tietze extension theorem.) *For any closed set $A \subset S$ and continuous function $g : A \to [a,b]$, there is a continuous extension $G : S \to [a,b]$ of g.*

2. *A Hausdorff space is paracompact iff it admits a C^0 partition of unity subordinate to any open covering.*

It is clear that if $\{(U_i, g_i)\}$ is a continuous partition of unity subordinate to the given open covering $\{V_\alpha\}$, then by definition $\{U_i\}$ is an open locally finite refinement. The converse—the existence of partitions of unity—is the hard part; the proof of this and of the equivalences of (i), (ii), and (iii) can be found for instance in Kelley [1975] and Choquet [1969, Section 6]. These results are important for the rich supply of continuous functions they provide. We shall not use these topological theorems in the rest of the book, but we do want their smooth versions on manifolds.

Note that if M is a manifold admitting partitions of unity subordinate to any open covering, then M is paracompact, and thus normal by Proposition 5.5.6. This already enables us to generalize (ii) and (iii) to the smooth (or C^k) situation.

5.5.8 Proposition. *Let M be a manifold admitting smooth (or C^k) partitions of unity. If A and B are closed disjoint sets then, there exists a smooth (or C^k) function $f : M \to [0,1]$ such that $f(A) = 0$ and $f(B) = 1$.*

Proof. As we saw, the condition on M implies that M is normal and thus there is an atlas $\{(U_\alpha, \varphi_\alpha)\}$ such that $U_\alpha \cap A \neq \varnothing$ implies $U_\alpha \cap B = \varnothing$. Let $\{(V_i, g_i)\}$ be a subordinate C^k partition on unity and $f = \Sigma g_i$, where the sum is over those i for which $V_i \cap B \neq \varnothing$. Then f is C^k, is one on B, and zero on A. ■

5.5.9 Theorem (Smooth Tietze Extension Theorem)**.** *Let M be a manifold admitting partitions of unity, and let $\pi : E \to M$ be a vector bundle with base M. Suppose $\sigma : A \to E$ is a C^k section defined on the closed set A (i.e., every point $a \in A$ has a neighborhood U_a and a C^k section $\sigma_a : U_a \to E$ extending σ). Then σ can be extended to a C^k global section $\Sigma : M \to E$. In particular, if $g : A \to \mathbf{F}$ is a C^k function defined on the closed set A, where $\mathbf{F}$ is a Banach space, then there is a C^k extension $G : M \to \mathbf{F}$; if g is bounded by a constant R, that is, $\|g(a)\| \leq R$ for all $a \in A$, then so is G.*

Proof. Consider the open covering $\{\, U_\alpha, M \backslash A \mid a \in A \,\}$ of M, with U_α given by the definition of smoothness on the closed set A. Let $\{(U_i, g_i)\}$ be a C^k partition of unity subordinate to this open covering and define $\sigma_i : U_i \to E$, by $\sigma_i = \sigma_a | U_i$ for all U_i and $\sigma_i \equiv 0$ on all U_i disjoint from U_a, $a \in A$. Then $g_i \sigma_i : U_i \to E$

is a C^k section on U_i and since $\operatorname{supp}(g_i\sigma_i) \subset \operatorname{supp}(g_i) \subset U_i$, it can be extended in a C^k manner to M by putting it equal to zero on $M \backslash U_i$. Thus $g_i\sigma_i : M \to E$ is a C^k-section of $\pi : E \to M$ and hence $\Sigma = \Sigma_i g_i \sigma_i$ is a C^k section; note that the sum is finite in a neighborhood of every point $m \in M$. Finally, if $a \in A$

$$\Sigma(a) = \sum_i g_i(a)\sigma_i(a) = \left(\sum_i g_i(a)\right)\sigma(a) = \sigma(a),$$

that is, $\Sigma | A = \sigma$.

The second part of the theorem is a particular case of the one just proved by considering the trivial bundle $M \times \mathbf{F} \to M$ and the section σ defined by $\sigma(m) = (m, g(m))$. The boundedness statement follows from the given construction, since all the g_i have values in $[0, 1]$. ■

Before discussing general questions on the existence of partitions of unity on Banach manifolds, we discuss the existence of Riemannian metrics. Recall that a Riemannian metric on a Hausdorff manifold M is a tensor field $g \in \mathcal{T}^0_2(M)$ such that for all $m \in M$, $g(m)$ is symmetric and positive definite. Our goal is to find topological conditions on an n-manifold that are necessary and sufficient to ensure the existence of Riemannian metrics. The proof of the necessary conditions will be simplified by first showing that any Riemannian manifold is a metric space. For this, define for $m, n \in M$,

$$\begin{aligned} d(m,n) = \inf\{\, \ell(\gamma) \mid \gamma : [0,1] \to M \text{ is a continuous piecewise} \\ C^1 \text{ curve with } \gamma(0) = m, \gamma(1) = n \,\}. \end{aligned}$$

Here $\ell(\gamma)$ is the ***length*** of the curve γ, defined by

$$\ell(\gamma) = \int_0^1 \|\dot{\gamma}(t)\| \, dt,$$

where $\dot{\gamma}(t) = d\gamma/dt$ is the tangent vector at $\gamma(t)$ to the curve γ and

$$\|\dot{\gamma}\| = [g_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))]^{1/2}$$

is its length.

5.5.10 Proposition. *d is a metric on each connected component of M whose metric topology is the original topology of M. If d is a complete metric, M is called a* ***complete Riemannian manifold****.*

Proof. Clearly $d(m, m) = 0$, $d(m, n) = d(n, m)$, and

$$d(m, p) \le d(m, n) + d(n, p),$$

by using the definition. Next we will verify that $d(m, n) > 0$ whenever $m \neq n$.

Let $m \in U \subset M$ where (U, φ) is a chart for M and suppose $\varphi(U) = U' \subset \mathbf{E}$. Then for any $u \in U$, $g(u)(v, v)^{1/2}$, defined for $v \in T_uM$, is a norm on T_uM. This is equivalent to the norm on $\mathbf{E}$, under the linear isomorphism $T_m\varphi$. Thus, if g' is the local expression for g, then $g'(u')$ defines an inner product on $\mathbf{E}$, yielding equivalent norms for all $u' \in U'$. Using continuity of g and choosing U' to be an open disk in $\mathbf{E}$, we can conclude that the norms $g'(u')^{1/2}$ and $g'(m)^{1/2}$, where $m' = \varphi(m)$ satisfy:

$$ag'\left(m'\right)^{1/2} \le g'\left(u'\right)^{1/2} \le bg'\left(m'\right)^{1/2}$$

for all $u' \in U'$, where a and b are positive constants. Thus, if $\eta : [0,1] \to U'$ is a continuous piecewise C^1 curve, then

$$\begin{aligned}\ell(\eta) &= \int_0^1 g'(\eta(t))(\dot\eta(t),\dot\eta(t))^{1/2}dt\\ &\geq a\int_0^1 g'(m')(\dot\eta(t),\dot\eta(t))^{1/2}dt\\ &\geq ag'(m')\left(\int_0^1 \dot\eta(t)dt, \int_0^1 \dot\eta(t)dt\right)^{1/2}\\ &\geq ag'(m')(\eta(1)-\eta(0),\eta(1)-\eta(0))^{1/2}.\end{aligned}$$

Here we have used the following property of the Bochner integral:

$$\left\|\int_a^b f(t)\,dt\right\| \leq \int_a^b \|f(t)\|dt,$$

valid for any norm on $\mathbf{E}$ (see the remarks following Definition 2.2.7).

Now let $\gamma : [0,1] \to M$ be a continuous piecewise C^1 curve, $\gamma(0) = m$, $\gamma(1) = n$, $m \in U$, where (U,φ), $\varphi : U \to U' \subset \mathbf{E}$ a chart of M, $\varphi(m) = 0$. If γ lies entirely in U, then $\varphi\circ\gamma = \eta$ lies entirely in U' and the previous estimate gives

$$\ell(\gamma) \geq ag'(m')(n',n')^{1/2} \geq ar,$$

where r is the radius of the disk U' in $\mathbf{E}$ about the origin and, $n' = \varphi(n)$. If γ is not entirely contained in U, then let r be the radius of a disk about the origin and let $c \in\,]0,1[$ be the smallest number for which $\gamma(c) \cap \varphi^{-1}(\{\, x \in \mathbf{E} \mid \|x\| = r\,\}) \neq \varnothing$. Then

$$\ell(\gamma) \geq \ell(\gamma|[0,c]) \geq ag(m')((\varphi\circ\gamma)(c),(\varphi\circ\gamma)(c))^{1/2} \geq ar.$$

Thus, we conclude $d(m,n) \geq ar > 0$.

The equivalence of the original topology of M and of the metric topology defined by d is clear if one notices that they are equivalent in every chart domain U, which in turn is implied by their equivalence in $\varphi(U)$. ■

Notice that the preceding proposition holds in infinite dimensions.

5.5.11 Proposition. *A connected Hausdorff n-manifold admits a Riemannian metric if and only if it is second countable. Hence for Hausdorff n-manifolds (not necessarily connected) paracompactness and metrizability are equivalent.*

Proof. If M is second countable, it admits partitions of unity by Theorem 5.5.7. Then the patching construction (Theorem 5.5.3) gives a Riemannian metric on M by choosing in every chart the standard inner product in $\mathbb{R}^n$.

Conversely, assume M is Riemannian. By Proposition 5.5.10 it is a metric space, which is locally compact and first countable, being locally homeomorphic to $\mathbb{R}^n$. By Theorem 1.6.14, it is second countable. ■

The main theorem on the existence of partitions of unity in the general case is as follows.

5.5.12 Theorem. *Any second-countable or paracompact manifold modeled on a separable Banach space with a C^k norm away from the origin admits C^k partitions of unity. In particular paracompact (or second countable) manifolds modeled on separable Hilbert spaces admit C^∞ partitions of unity.*

5.5.13 Corollary. *Paracompact (or second countable) Hausdorff manifolds modeled on separable real Hilbert spaces admit Riemannian metrics.*

Theorem 5.5.12 will be proved in Supplements 5.5A and 5.5B.

There are Hausdorff nonparacompact n-manifolds. These manifolds are necessarily nonmetrizable and do not admit partitions of unity. The standard example of a one-dimensional nonparacompact Hausdorff manifold is the "long line." In dimensions 2 and 3 such manifolds are constructed from the Prüfer manifolds. Since nonparacompact manifolds occur rarely in applications, we refer the reader to Spivak [1979, Volume 1, Appendix A] for the aforementioned examples.

Partitions of unity are an important technical tool in many proofs. We illustrate this with the sample theorem below which combines differential topological ideas of §3.5, the local and global existence and uniqueness theorem for solutions of vector fields, and partitions of unity. More applications of this sort can be found in the exercises.

5.5.14 Theorem (Ehresmann Fibration Theorem). *A proper submersion $f : M \to N$ of finite dimensional manifolds with M paracompact is a locally trivial fibration, that is, for any $p \in N$ there exists an open neighborhood V of u in N and a diffeomorphism $\varphi : V \times f^{-1}(p) \to f^{-1}(V)$ such that $f(\varphi(x,u)) = x$ for all $x \in V$ and all $u \in f^{-1}(p)$.*

Proof. Since the statement is local we can replace M, N by chart domains and, in particular, we can assume that $N = \mathbb{R}^n$ and $p = 0 \in \mathbb{R}^n$. We claim that there are smooth vector fields $X_1, \ldots, X_n$ on M such that X_i is f-related to $\partial/\partial x^i \in \mathfrak{X}(\mathbb{R}^n)$. Indeed, around any point in M such vector fields are easy to obtain using the implicit function theorem (see Theorem 3.5.2). Cover M with such charts, choose a partition of unity subordinate to this covering, and patch these vector fields by means of this partition of unity to obtain $X_1, \ldots, X_n$, f-related to $\partial/\partial x^1, \ldots, \partial/\partial x^n$, respectively.

Let $F^k_{t(k)}$ denote the flow of X_k with time variable $t(k)$, $k = 1, \ldots, n$ and let $\mathbf{t} = (t(1), \ldots, t(n)) \in \mathbb{R}^n$. If $\|\mathbf{t}\| < C$, then the integral curves of each X_k starting in $f^{-1}(\{\, \mathbf{u} \in \mathbb{R}^n \mid \|\mathbf{u}\| \leq C \,\})$ stay in $f^{-1}\{\, \mathbf{v} \in \mathbb{R}^n \mid \|\mathbf{v}\| \leq 2C \,\}$, since by Proposition 4.2.4

$$(f \circ F^k_{t(k)})(\mathbf{y}) = (f^1(\mathbf{y}), \ldots, f^k(\mathbf{y}) + t(k), \ldots, f^n(\mathbf{y})). \tag{5.5.1}$$

Therefore, since f is proper, Proposition 4.1.19 implies that the vector fields $X_1, \ldots, X_n$ are complete.

Finally, let $\varphi : \mathbb{R}^n \times f^{-1}(0) \to M$ be given by

$$\varphi(t(1), \ldots, t(n), u) = (F^1_{t(1)} \circ \cdots \circ F^n_{t(n)})(u)$$

and note that φ is smooth (see Proposition 4.1.17). The map $\varphi^{-1} : M \to \mathbb{R}^n \times f^{-1}(0)$ given by

$$\varphi^{-1}(m) = (f(m), (F^n_{-t(n)} \circ \cdots \circ F^1_{-t(1)})(m))$$

is smooth and is easily checked to be the inverse of φ. Finally, $(f \circ \varphi)(\mathbf{t}, u) = \mathbf{t}$ by equation (5.5.1) since $f(u) = 0$. ■

Supplement 5.5A
Partitions of Unity: Reduction to the Local Case

We begin with some topological preliminaries. Let S be a paracompact space. If $\{U_\beta\}$ is an open covering of S, it can be refined to a locally finite covering $\{W_\beta\}$. The first lemma below will show that we can shrink this covering further to get another one, $\{V_\alpha\}$ such that $\mathrm{cl}(V_\alpha) \subset W_\alpha$ with the same indexing set.

A technical device used in the proof is the concept of a ***well-ordered*** set. An ordered set A in which any two elements can be compared is called well-ordered if every subset has a smallest element (see the introduction to Chapter 1).

5.5.15 Lemma (Shrinking Lemma). *Let S be a normal space and let $\{W_\alpha\}_{\alpha\in A}$ be a locally finite open covering of S. Then there is a locally finite open refinement $\{V_\alpha\}_{\alpha\in A}$ (with the same indexing set) such that* $\mathrm{cl}(V_\alpha) \subset W_\alpha$.

Proof. Well-order the indexing set A and call its smallest element $\alpha(0)$. The set C_0 defined as $C_0 = S\backslash\bigcup_{\alpha>\alpha(0)} W_\alpha$ is closed, so by normality there exists an open set $V_{\alpha(0)}$ such that $C_0 \subset \mathrm{cl}(V_{\alpha(0)}) \subset W_{\alpha(0)}$. If V_γ is defined for all $\gamma < \alpha$, put

$$C_\alpha = S\backslash\{(\bigcup_{\gamma<\alpha} V_\gamma) \cup (\bigcup_{\gamma>\alpha} W_\gamma)\}$$

and by normality find V_α such that $C_\alpha \subset \mathrm{cl}(V_\alpha) \subset W_\alpha$. The collection $\{V_\alpha\}_{\alpha\in A}$ is the desired locally finite refinement of $\{W_\alpha\}_{\alpha\in A}$, provided we can show that it covers S. Given $s \in S$, by local finiteness of the covering $\{W_\alpha\}_{\alpha\in A}$, s belongs to only a finite collection of them, say $W_1, W_2, \ldots, W_n$, corresponding to the elements $\alpha_1, \ldots, \alpha_n$ of the index set. If β denotes the maximum of the elements $\alpha_1, \ldots, \alpha_n$, then $s \notin W_\gamma$ for all $\gamma > \beta$, so that if in addition $s \notin V_\gamma$ for all $\gamma < \beta$, then $s \in C_\beta \subset V_\beta$, that is, $s \in V_\beta$. ■

5.5.16 Lemma (A. H. Stone). *Every pseudometric space is paracompact.*

Proof. Let $\{U_\alpha\}_{\alpha\in A}$ be an open covering of the pseudometric space S with distance function d. Put

$$U_{n,\alpha} = \{\, x \in U_\alpha \mid d(x, S\backslash U_\alpha) \geq 1/2^n \,\}.$$

By the triangle inequality, we have the inequality

$$d(U_{n,\alpha}, S\backslash U_{n+1,\alpha}) \geq \frac{1}{2^n} - \frac{1}{2^{n+1}} = \frac{1}{2^{n+1}}.$$

Well-order the indexing set A and let $V_{n,\alpha} = U_{n,\alpha}\backslash\bigcup_{\beta<\alpha} U_{n+1,\beta}$. If $\gamma, \delta \in A$, we have $V_{n,\gamma} \subset S\backslash U_{n+1,\delta}$, if $\gamma < \delta$, and $V_{n,\delta} \subset S\backslash U_{n+1,\gamma}$ if $\delta < \gamma$. But in both cases we have $d(V_{n,\gamma}, V_{n,\delta}) \geq 1/2^{n+1}$. Define

$$W_{n,\alpha} = \{\, s \in S \mid d(s, V_{n,\alpha}) < 1/2^{n+3} \,\},$$

and observe that $d(W_{n,\alpha}, W_{n,\beta}) \geq 1/2^{n+2}$. Thus for a fixed n, every point $s \in S$ has a neighborhood intersecting at most one member of the family $\{\, W_{n,\alpha} \mid \alpha \in A \,\}$. Hence $\{\, W_{n,\alpha} \mid n \in \mathbb{N}, \alpha \in A \,\}$ is a locally finite open refinement of $\{U_\alpha\}$. ■

Let us now turn to the question of the existence of partitions of unity subordinate to any open covering.

5.5.17 Proposition (R. Palais). *Let M be a paracompact manifold modeled on the Banach space* $\mathbf{E}$. *The following are equivalent:*

(i) *M admits C^k partitions of unity;*

(ii) *any open covering of M admits a locally finite refinement by C^k carriers;*

(iii) *for any open sets O_1, O_2 such that $\mathrm{cl}(O_1) \subset O_2$, there exists a C^k carrier V such that $O_1 \subset V \subset O_2$;*

(iv) *every chart domain of M admits C^k partitions of unity subordinate to any open covering;*

(v) $\mathbf{E}$ *admits C^k partitions of unity subordinate to any open covering of* $\mathbf{E}$.

Proof. (i) ⇒ (ii). If $\{(U_i, g_i)\}$ is a C^k-partition of unity subordinate to an open covering, then clearly $\mathrm{carr}\ g_i$ forms a locally finite refinement of the covering by C^k carriers.

(ii) ⇒ (iii). Let $\{V_\alpha\}_{\alpha\in A}$ be a locally finite refinement of the open covering $\{O_2, S\backslash\mathrm{cl}(O_1)\}$ by C^k carriers and denote by $f_\alpha \in \mathcal{F}^k(M)$, the function for which $\mathrm{carr}\, f_\alpha = V_\alpha$. Let

$$B = \{\, \alpha \in A \mid V_\alpha \subset O_2 \,\}.$$

Put $V = \bigcap_{\beta \in B} V_\beta$, $f = \Sigma_{\beta \in B} f_\beta$ and notice that $O_1 \subset V \subset O_2$, $\mathrm{carr}(f) = V$.

(iii) $\Rightarrow$ (iv). Let U be any chart domain of M. Then U is diffeomorphic to an open set in $\mathbf{E}$ which is a metric space, so is paracompact by Stone's theorem (Lemma 5.5.16). Let $\{U_\alpha\}$ be an arbitrary open covering of U and $\{V_\beta\}$ be a locally finite refinement. By the shrinking lemma we may assume that $\mathrm{cl}(V_\beta) \subset U$. Again by the shrinking lemma, refine further to a locally finite covering $\{W_\beta\}$ such that $\mathrm{cl}(W_\beta) \subset V_\beta$. But by (iii) there exists a C^k-carrier O_β such that $W_\beta \subset O_\beta \subset V_\beta$, and so $\{O_\beta\}$ is a locally finite refinement of $\{U_\alpha\}$ by C^k-carriers, whose corresponding functions we denote by f_β. Thus $f = \Sigma_\beta f_\beta$ is a C^k map and $\{(V_\beta, f_\beta/f)\}$ is a C^k partition of unity subordinate to $\{U_\alpha\}$.

(iv) $\Rightarrow$ (v). Consider now any open covering $\{U_\alpha\}_{\alpha \in A}$ of $\mathbf{E}$ and let (U, φ) be an arbitrary chart of M. Refine first the covering of $\mathbf{E}$ by taking the intersections of all its elements with all translates of $\varphi(U)$. Since $\mathbf{E}$ is paracompact, refine again to a locally finite open covering $\{V_\beta\}$. The inverse images by translations and φ of these open sets are subsets of U, hence chart domains, and thus by (iv) they admit partitions of unity subordinate to any covering. Thus every V_β admits a C^k partition of unity subordinate to any open covering, for example to $\{\, V_\beta \cap U_\alpha \mid \alpha \in A \,\}$; call it $\{g_i^\beta\}$. Then $g = \Sigma_{i,\beta} g_i^\beta$ is a C^k map and the double-indexed set of functions g_i^β / g forms a C^k partition of unity of $\mathbf{E}$.

(v) $\Rightarrow$ (iv). If $\mathbf{E}$ admits C^k partitions of unity subordinate to any open covering, then so does every open subset by the (already proved) implication (i) $\Rightarrow$ (ii) applied to $M = \mathbf{E}$, which is paracompact by Theorem 5.5.16. Thus if (U, φ) is a chart on M, U admits partitions of unity, since $\varphi(U)$ does.

Finally, we show (iv) implies (i) . Choosing a locally finite atlas, this proof repeats the one given in the last part of (iv) $\Rightarrow$ (v). ■

As an application of this proposition we get the following.

5.5.18 Proposition. *Every paracompact n-manifold admits C^∞-partition of unity.*

Proof. By Proposition 5.5.17(ii) and (v) it suffices to show that every open set in $\mathbb{R}^n$ is a C^∞ carrier. Any open set U is a countable union of open disks D_i. By Lemma 4.2.13, $D_i = \mathrm{carr}\,(f_i)$, for some C^∞ function $f_i : \mathbb{R}^n \to \mathbb{R}$. Put

$$M_i = \sup\{\, \|D^k f_i(x)\| \mid x \in \mathbb{R}^n,\ k \leq i \,\}$$

and let

$$f = \sum_{i=1}^{\infty} \frac{f_i}{2^i M_i}.$$

By Exercise 2.4-10, f is a C^∞ function for which $\mathrm{carr}(f) = U$ clearly holds. ■

In particular, second-countable n-manifolds admit partitions of unity, recovering Theorem 5.5.7.

SUPPLEMENT 5.5B

Partitions of Unity: The Local Case

Proposition 5.5.17 reduces the problem of the existence of partitions of unity to the local one, namely finding partitions of unity in Banach spaces. This problem has been studied by Bonic and Frampton [1966] for separable Banach spaces.

5.5.19 Proposition (Bonic and Frampton [1966]). *Let* $\mathbf{E}$ *be a separable Banach space. The following are equivalent.*

(i) *Any open set of* $\mathbf{E}$ *is a* C^k *carrier.*

(ii) $\mathbf{E}$ *admits* C^k *partitions of unity subordinate to any open covering of* $\mathbf{E}$.

(iii) *There exists a bounded nonempty* C^k *carrier in* $\mathbf{E}$.

Proof. By Proposition 5.5.17, (i) and (ii) are equivalent since $\mathbf{E}$ is paracompact by Lemma 5.5.16. It remains to be shown that (iii) implies (i), since clearly (ii) implies (iii).

This proceeds in several steps. First, we show that any neighborhood contains a C^k carrier. Let U be any open set and let $\operatorname{carr}(f) \subset D_r(0)$ be the bounded carrier given by (iii), $f \in C^k(\mathbf{E})$, $f \geq 0$. Let $e \in U$, fix $e_0 \in \operatorname{carr}(f)$, and choose $\varepsilon > 0$ such that $D_\varepsilon(e) \subset U$. Define $g \in C^k(\mathbf{E})$, $g \geq 0$ by

$$g(v) = f(K(v-e)+e_0), \quad K > 0,$$

where K remains to be determined from the condition that $\operatorname{carr}(g) \subset D_\varepsilon(e)$. An easy computation shows that if $K > (r + \|e_0\|)/\varepsilon$, this inclusion is verified. Since $e \in \operatorname{carr}(g), \operatorname{carr}(g)$ is an open neighborhood of e.

Second, we show that any open set can be covered by a countable locally finite family of C^k carriers. By the first step, the open set U can be covered by a family of C^k carriers. By Lindelöf's lemma 1.1.6, $U = \bigcup_n V_n$ where V_n is a C^k carrier, the union being over the positive integers. We need to find a refinement of this covering by C^k carriers. Let $f_n \in C^k(\mathbf{E})$ be such that $\operatorname{carr} f_n = V_n$. Define

$$U_n = \{\, e \in \mathbf{E} \mid f_n(e) > 0,\ f_i(e) < 1/n \text{ for all } i < n \,\}.$$

Clearly $U_1 = V_1$ and inductively

$$U_n = V_n \cap \left[\bigcap_{i<n} f_i^{-1}(]-\infty, 1/n[) \right].$$

By the composite mapping theorem, the inverse image of a C^k carrier is a C^k carrier, so that $f_i^{-1}(]-\infty, 1/n[)$ is a C^k carrier, since $]\infty, 1/n[$ is a C^k carrier in $\mathbb{R}$ (see the proof of Proposition 5.5.18). Finite intersections of C^k carriers is a C^k carrier (just take the product of the functions in question) so that U_n is also a C^k carrier. Clearly $U_n \subset V_n$. We shall prove that $\{U_n\}$ is a locally finite open covering of U. Let $e \in U$. If $e \in V_n$ for all n, then clearly $e \in U_1 = V_1$. If not, then there exists a smallest n, say N, such that $e \in V_N$. Then $f_i(e) = 0$ for $i < N$ and thus

$$e \in U_N = \{\, e \in \mathbf{E} \mid f_N(e) > 0,\ f_i(e) < 1/N \text{ for all } i < N \,\}.$$

Thus, the sets U_n cover U. This open covering is also locally finite for if $e \in V_n$ and N is such that $f(e) > 1/N$, then the neighborhood $\{\, u \in U \mid f_n(e) > 1/N \,\}$ has empty intersections with all U_m for $m > N$.

Third we show that the open set U is a C^k-carrier. By the second step, $U = \bigcup_n U_n$, with U_n a locally finite open covering of U by C^k carriers. Then $f = \Sigma_n f_n$ is C^k, $f(e) \geq 0$ for all $e \in \mathbf{E}$ and $\operatorname{carr}(f) = U$. ■

The separability assumption was used only in showing that (iii) implies (i). There is no general theorem known to us for nonseparable Banach spaces. Also, it is not known in general whether Banach spaces admit bounded C^k carriers, for $k \geq 1$. However, we have the following.

5.5.20 Proposition. *If the Banach space* $\mathbf{E}$ *has a norm* C^k *away from its origin,* $k \geq 1$*, then* $\mathbf{E}$ *has bounded* C^k*-carriers.*

Proof. By Corollary 4.2.14 there exists $\varphi : \mathbb{R} \to \mathbb{R}$, C^∞ with compact support and equal to one in a neighborhood of the origin. If $\|\cdot\| : \mathbf{E}\backslash\{0\} \to \mathbb{R}$ is C^k, $k \geq 1$, then $\varphi \circ \|\cdot\| : \mathbf{E}\backslash\{0\} \to \mathbb{R}$ is a nonzero map which is C^k, has bounded support $\|\cdot\|^{-1}(\operatorname{supp}\varphi)$, and can be extended in a C^k manner to $\mathbf{E}$. ■

Theorem 5.5.12 now follows from Propositions 5.5.20, 5.5.19, and 5.5.17.

The situation with regard to Banach subspaces and submanifolds is clarified in the following proposition, whose proof is an immediate consequence of Propositions 5.5.19 and 5.5.17.

5.5.21 Proposition.

(i) *If* $\mathbf{E}$ *is a Banach space admitting* C^k *partitions of unity then so does any closed subspace,*

(ii) *If a manifold admits* C^k *partitions of unity subordinate to any open covering, then so does any submanifold.*

We shall not develop this discussion of partitions of unity on Banach manifolds any further, but we shall end by quoting a few theorems that show how intimately connected partitions of unity are with the topology of the model space. By Propositions 5.5.19 and 5.5.20, for separable Banach spaces one is interested whether the norm is C^k away from the origin. Restrepo [1964] has shown that a separable Banach space has a C^1 norm away from the origin if and only if its dual is separable. Bonic and Reis [1966] and Sundaresan [1967] have shown that if the norms on $\mathbf{E}$ and $\mathbf{E}^*$ are differentiable on $\mathbf{E}\backslash\{0\}$ and $\mathbf{E}^*\backslash\{0\}$, respectively, then $\mathbf{E}$ is reflexive, for $\mathbf{E}$ a real Banach space (not necessarily separable). Moreover, $\mathbf{E}$ is a Hilbert space if and only if the norms on $\mathbf{E}$ *and* $\mathbf{E}^*$ are twice differentiable away from the origin. This result has been strengthened by Leonard and Sunderesan [1973], who show that a real Banach space is isometric to a Hilbert space if and only if the norm is C^2 away from the origin and the second derivative of $e \mapsto \|e\|^2/2$ is bounded by 1 on the unit sphere; see Rao [1972] for a related result. Palais [1965b] has shown that any paracompact Banach manifold admits Lipschitz partitions of unity.

Because of the importance of the differentiability class of the norm in Banach spaces there has been considerable work in the direction of determining the exact differentiability class of concrete function spaces. Thus Bonic and Frampton [1966] have shown that the canonical norms on the spaces $L^p(\mathbb{R})$, $\ell^p(\mathbb{R})$, $p \geq 1$, $p < \infty$ are C^∞ away from the origin if p is even, C^{p-1} with $\mathbf{D}(\|\cdot\|^{p-1})$ Lipschitz, if p is odd, and $C^{[p]}$ with $\mathbf{D}^{[p]}(\|\cdot\|^p)$ Hölder continuous of order $p-[p]$, if p is not an integer. The space c_0 of sequences of real numbers convergent to zero has an equivalent norm that is C^∞ away from the origin, a result due to Kuiper. Using this result, Frampton and Tromba [1972] show that the Λ-spaces (closures of C^∞ in the Hölder norm) admit a C^∞ norm away from the origin. The standard norm on the Banach space of continuous real valued functions on $[0,1]$ is nowhere differentiable. Moreover, since $C^0([0,1],\mathbb{R})$ is separable with nonseparable dual, it is impossible to find an equivalent norm that is differentiable away from the origin. To our knowledge it is still an open problem whether $C^0([0,1],\mathbb{R})$ admits C^∞ partitions of unity for $k \geq 1$.

Finally, the only results known to us for nonseparable Hilbert spaces are those of Wells [1971, 1973], who has proved that nonseparable Hilbert space admits C^2 partitions of unity. The techniques used in the proof, however, do not seem to indicate a general way to approach this problem.

Supplement 5.5C

Simple Connectivity of Fiber Bundles

The goal of this supplement is to discuss the homotopy lifting property for locally trivial continuous fiber bundles over a paracompact base. This theorem is shown to imply an important criterion on the simple connectedness of the total space of fiber bundles with paracompact base.

5.5.22 Theorem (Homotopy Lifting Theorem)**.** *Let* $\pi : E \to B$ *be a locally trivial* C^0 *fiber bundle and let* M *be a paracompact topological space. If* $h : [0,1] \times M \to B$ *is a continuous homotopy and* $f : M \to E$ *is any continuous map satisfying* $\pi \circ f = h(0,\cdot)$*, there exists a continuous homotopy* $H : [0,1] \times M \to E$

satisfying $\pi \circ H = h$ and $H(0, \cdot) = f$. If in addition, h fixes some point $m \in M$, that is $h(t, m)$ is constant for t in a segment Δ of $[0,1]$, then $H(t, m)$ is also constant for $t \in \Delta$.

Remark. The property in the statement of the theorem is called the ***homotopy lifting property***. A ***Hurewicz fibration*** is a continuous surjective map $\pi : E \to B$ satisfying the homotopy lifting property relative to any topological space M. Thus the theorem above says that a *locally trivial C^0 fiber bundle is a Hurewicz fibration relative to paracompact spaces*. ♦

See Steenrod [1957] and Huebsch [1955] for the proof.

5.5.23 Corollary. *Let $\pi : E \to B$ be a C^0 locally trivial fiber bundle. If the base B and the fiber F are simply connected, then E is simply connected.*

Proof. Let $c : [0,1] \to E$ be a loop, $c(0) = c(1) = e_0$. Then $d = \pi \circ c$ is a loop in B based at $\pi(e_0) = b_0$. Since B is simply connected there is a homotopy $h : [0,1] \times [0,1] \to B$ such that $h(0,t) = d(t)$, $h(1,t) = b_0$ for all $t \in [0,1]$, and $h(s,0) = h(s,1) = b_0$ for all $s \in [0,1]$. By the homotopy lifting theorem there is a homotopy $H : [0,1] \times [0,1] \to E$ such that $\pi \circ H = h$, $H(0, \cdot) = c$, and

$$H(s,0) = H(0,0) = c(0) = e_0, \quad H(s,1) = c(1) = e_0.$$

Since $(\pi \circ H)(1,t) = h(1,t) = b_0$, it follows that $t \mapsto H(1,t)$ is a path in $\pi^{-1}(b_0)$ starting at $H(1,0) = e_0$ and ending also at $H(1,1) = e_0$. Since $\pi^{-1}(b_0)$ is simply connected, there is a homotopy $k : [1,2] \times [0,1] \to \pi^{-1}(b_0)$ such that

$$k(1,t) = H(1,t), \quad k(2,t) = e_0$$

for all $t \in [0,1]$ and

$$k(s,0) = k(s,1) = e_0$$

for all $s \in [1,2]$. Define the continuous homotopy $K : [0,2] \times [0,1] \to E$ by

$$K(s,t) = \begin{cases} H(s,t), & \text{if } s \in [0,1]; \\ k(s,t), & \text{if } s \in [1,2], \end{cases}$$

and note that

$$K(0,t) = H(0,t) = c(t), \quad K(2,t) = k(2,t) = e_0$$

for any $t \in [0,1]$, and $K(s,1) = e_0$ for any $s \in [0,2]$. Thus, c is contractible to e_0 and E is therefore simply connected. ■

Exercises

◇ **5.5-1** (Whitney)**.** Show that any closed set F in $\mathbb{R}^n$ is the inverse image of 0 by a C^∞ real-valued positive function on $\mathbb{R}^n$. Generalize this to any n-manifold.
HINT: Cover $\mathbb{R}^n \backslash F$ with a sequence of open disks D_n and choose for each n a smooth function $\chi_n \geq 0$, satisfying $\chi_n | D_n > 0$, with the absolute value of χ_n and all its derivatives $\leq 2^n$. Set $\chi = \Sigma_{n \geq 0} \chi_n$.

◇ **5.5-2.** In a paracompact topological space, an open subset need not be paracompact. Prove the following.

(i) If every open subset of a paracompact space is paracompact, then any subspace is paracompact.

(ii) Every open submanifold of a paracompact manifold is paracompact.

HINT: Use chart domains to conclude metrizability.

◇ **5.5-3.** Let $\pi : E \to M$ be a vector bundle, $E' \subset E$ a subbundle and assume M admits C^k partitions of unity subordinate to any open covering. Show that E' splits in E, that is, there exists a subbundle E'' such that $E = E'' \oplus E''$.
HINT: The result is trivial for local bundles. Construct for every element of a locally finite covering $\{U_i\}$ a vector bundle map f_i whose kernel is the complement of $E'|U_i$. For $\{(U_i, g_i)\}$ a C^k partition of unity, put $f = \Sigma_i g_i f_i$ and show that $E = E' \oplus \ker f$.

◇ **5.5-4.** Let $\pi : E \to M$ be a vector bundle over the base M that admits C^k partitions of unity and with the fibers of E modeled on a Hilbert space. Show that E admits a C^k bundle metric, that is, a C^k map $g : M \to T^0_2(E)$ that is symmetric, strongly nondegenerate, and positive definite at every point $m \in M$.

◇ **5.5-5.** Let $E \to M$ be a line bundle over the manifold M admitting C^k partitions of unity sub-ordinate to any open covering. Show that $E \times E$ is trivial
HINT: $E \times E = L(E^*, E)$ and construct a local base that can be extended.

◇ **5.5-6.** Assume M admits C^k partitions of unity. Show that any submanifold of M diffeomorphic to S^1 is the integral curve of a C^k vector field on M.

◇ **5.5-7.** Let M be a connected paracompact manifold. Show that there exists a C^∞ proper mapping $f : M \to \mathbb{R}^k$.
HINT: M is second countable, being Riemannian. Show the statement for $k = 1$, where $f = \Sigma_{i \geq 1} i\varphi_i$, and $\{\varphi_i\}$ is a countable partition of unity.

◇ **5.5-8.** Let M be a connected paracompact n-manifold and $X \in \mathfrak{X}(M)$. Show that there exist $h \in \mathcal{F}(M), h > 0$ such that $Y = hX$ is complete.
HINT: With f as in Exercise 5.5-7 put $h = \exp\{-(X[f])^2\}$ so that $|Y[f]| \leq 1$. Hence $(f \circ c)(]a, b[)$ is bounded for any integral curve c of Y and $]a, b[$ in the domain of c.

◇ **5.5-9.** Let M be a paracompact, non-compact manifold.

(i) Show that there exists a locally finite sequence of open sets $\{ U_i \mid i \in \mathbb{Z} \}$ such that $U_i \cap U_{i+1} \neq \varnothing$ unless $j = i-1, i, i+1$, and each U_i is a chart domain diffeomorphic (by the chart map φ_i) with the open unit ball in the model space of M. See Figure 5.5.1.

HINT: Let $\mathcal{V}$ be a locally finite open cover of M with chart domains diffeomorphic by their chart maps to the open unit ball such that no finite subcover of $\mathcal{V}$ covers M, and no two elements of $\mathcal{V}$ include each other. Let U_0, U_1 be distinct elements of $\mathcal{V}$, $U_0 \cap U_1 \neq \varnothing$. Let $U_{-1} \in \mathcal{V}$ be such that $U_{-1} \cap U_0 \neq \varnothing$; $U_{-1} \cap U_1 = \varnothing$; such a U_{-1} exists by local finiteness of $\mathcal{V}$. Now use induction.

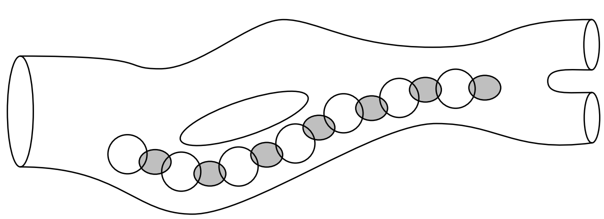

FIGURE 5.5.1. A chart chain

(ii) Use (i) to show that there exists an embedding of $\mathbb{R}$ in M as a closed manifold.

HINT: Let c_0 be a smooth curve in U_0 diffeomorphic by φ_0 to $]-1,1[$ connecting a point in $U_{-1} \cap U_0$ to a point in $U_0 \cap U_1$. Next, extend $c_0 \cap U_1$ smoothly to a curve c_1, diffeomorphic by φ_1 to $]0,2[$ ending inside $U_1 \cap U_2$; show that $c_1 \cap U_0$ extends the curve $c_0 \cap U_1$ inside $U_0 \cap U_1$. Now use induction.

(iii) Show that on each non-compact paracompact manifold admitting partitions of unity there exists a non-complete vector field.

HINT: Embed $\mathbb{R}$ in M by (ii) and on $\mathbb{R}$ consider the vector field $\dot{x} = x^2$. Extend it to M via partitions of unity.

◇ **5.5-10.** Show that every compact n-manifold embeds in some $\mathbb{R}^k$ for k big enough in the following way. If $\{(U_i, \varphi_i)\}_{i=1,\ldots,N}$ is a finite atlas with $\varphi_i(U_i)$ the ball of radius 2 in $\mathbb{R}^n$, let $\chi \in C^\infty(\mathbb{R}^n)$, $\chi = 1$ on the ball of radius 1 and $\chi = 0$ outside the ball of radius 2. Put $f_i = (\chi \circ \varphi_i) \cdot \varphi_i : M \to \mathbb{R}^n$, where $f_i = 0$ outside U_i. Show that f_i is C^∞ and that $\psi : M \to \mathbb{R}^{Nn} \times \mathbb{R}^N$, defined by

$$\psi(m) = (f_1(m), \ldots, f_N(m), \chi(\varphi_1(m)), \ldots, \chi(\varphi_N(m)))$$

is an embedding.

◇ **5.5-11.** Let g be a Riemannian metric on M.

(i) Show that if N is a submanifold, its g-normal bundle $\nu_g(N) = \{\, v \in T_n M \mid n \in N,\ v \perp T_n N \,\}$ is a subbundle of TM.

(ii) Show that $TM|N = \nu_g(N) \oplus TN$.

(iii) If h is another Riemannian metric on M, show that $\nu_g(N)$ is a vector bundle isomorphic to $\nu_h(N)$.

◇ **5.5-12.** Show that if $f : M \to N$ is a proper surjective submersion with M paracompact and N connected, then it is a locally trivial fiber bundle.
HINT: To show that all fibers are diffeomorphic, connect a fixed point of N with any other point by a smooth path and cover the path with the neighborhoods in N given by Theorem 5.5.14.