Contents

These notes are based mainly on the textbooks by Tu [4] and Boothby [1]. However, I frequently add my own proofs, examples, remarks, and results.

1 Smooth Functions on a Euclidean Space

1.1 C^∞ Versus Analytic Functions

Definition 1.1. Let k be a nonnegative integer and U an open subset of $\mathbb{R}^n$. A function $f : U \to \mathbb{R}$ is said to be C^k at p if its partial derivatives $\partial^j f/\partial x^{i_1} \cdots \partial x^{i_j}$ of all orders $j \leq k$ exist and are continuous at p. The function $f : U \to \mathbb{R}$ is C^∞ at p if it is C^k at p for all $k \geq 0$. We say that f is C^k (or C^∞) on U if it is C^k (or C^∞) at every point in U. The function f is called *real analytic* at p if in some neighborhood of p it is equal to its Taylor series at p, i.e.

$$f(x) = f(p) + \sum_i \frac{\partial f}{\partial x^i}(p)(x^i - p^i) + \frac{1}{2!}\sum_{i,j} \frac{\partial^2 f}{\partial x^i \partial x^j}(p)(x^i - p^i)(x^j - p^j) + \cdots .$$

Note that a real analytic function is C^∞, but not conversely.

1.2 Taylor's Theorem with Remainder

Definition 1.2. We say that a subset S of $\mathbb{R}^n$ is *star shaped* with respect to a point $p \in S$ if for every $x \in S$ the line segment from p to x lies in S.

Theorem 1.3. (Taylor's Theorem With Remainder) *Let f be a C^∞ function on an open subset U of $\mathbb{R}^n$ that is star shaped with respect to a point $p = (p^1, \ldots, p^n)$ in U. Then there are C^∞ functions $g_1(x), \ldots, g_n(x)$ on U such that*

$$f(x) = f(p) + \sum_1^n (x^i - p^i) g_i(x), \quad g_i(p) = \frac{\partial f}{\partial x^i}(p).$$

2 Tangent Vectors in $\mathbb{R}^n$ as Derivations

2.1 The Directional Derivative

Definition 2.1. For a point $p \in \mathbb{R}^n$, the *tangent space at* p, denoted $T_p(\mathbb{R}^n)$, is the set of all arrows whose base is at p. We may identify these arrows with points of $\mathbb{R}^n$ itself. Elements of $T_p(\mathbb{R}^n)$ are called *tangent vectors* (or just *vectors*) at p in $\mathbb{R}^n$. We write $p = (p^1, \dots, p^n)$ for the point p in $\mathbb{R}^n$, and we write

$$v = \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix} \quad \text{or} \quad v = \langle v^1, \dots, v^n \rangle$$

for a vector v in $T_p(\mathbb{R}^n)$.

Definition 2.2. Let $p \in \mathbb{R}^n$ and suppose f is a C^∞ function on a neighborhood of p. The *directional derivative* of f in the direction of v at p is defined to be

$$D_v f = \lim_{t \to 0} \frac{f(c(t)) - f(p)}{t} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(c(t)),$$

where $c : \mathbb{R} \to \mathbb{R}$ is the path defined by

$$c(t) = p + tv = (p^1 + tv^1, \dots, p^n + tv^n).$$

Although the letter p was suppressed in the notation D_v, it is understood that the directional derivative $D_v f$ is evaluated at the point p. Letting c^i denote the ith component function of $c(t)$, i.e. $c^i(t) = p^i + tv^i$, a simple application of the chain rule shows that

$$D_v f = \sum_1^n \frac{\mathrm{d}c^i}{\mathrm{d}t}(0) \frac{\partial f}{\partial x^i}(p) = \sum_1^n v^i \frac{\partial f}{\partial x^i}(p).$$

Thus we may also write D_v in operator notation as

$$D_v = \sum_1^n v^i \left.\frac{\partial}{\partial x^i}\right|_p .$$

2.2 Germs of Functions

Definition 2.3. Let $p \in \mathbb{R}^n$ and consider the set $\mathcal{S}$ of all pairs (f, U) where U is a neighborhood of p and $f : U \to \mathbb{R}$ is C^∞. Define an equivalence relation $\sim$ on $\mathcal{S}$ as follows. We say $(f, U) \sim (g, V)$ iff there exists an open set $W \subset U \cap V$ such that $f = g$ on W. The equivalence class containing (f, U) is called the *germ of f at p*. The quotient $S/\sim$ is the set of all germs of C^∞ functions on $\mathbb{R}^n$ at p, and it is denoted $C_p^\infty(\mathbb{R}^n)$, or just C_p^∞.

Note that C_p^∞ is a vector space over $\mathbb{R}$. (This requires careful definition of addition and scalar multiplication of germs. See the homework.)

Definition 2.4. Let K be a field. An *algebra* over K is a vector space V over K with a multiplication

$$\mu : V \times V \to V,$$

denoted $\mu(a, b) = a \times b$, such that for all $x, y, z \in V$ and $t \in K$ we have

1. (associativity) $(x \times y) \times z = x \times (y \times z)$,
2. (distributivity) $(x+y)\times z = x\times z+y\times z$ and $x\times(y+z) = x\times y+x\times z$.
3. (homogeneity) $t(x \times y) = (tx) \times y = x \times (ty)$.

In other words, an algebra over K is a ring V that's also a vector space over K such that the ring multiplication operation satisfies homogeneity condition 3 above.

2.3 Derivations at a Point

Definition 2.5. Let V, W be vector spaces over the field K. A map $T : V \to W$ is called *linear* (or *K-linear* for emphasis) if for all $c \in K$ and $x, y \in V$ we have

1. $T(x + y) = T(x) + T(y)$,
2. $T(cx) = cT(x)$.

Equivalently, if $T(cx + y) = cT(x) + T(y)$ for all $x, y \in V$ and $c \in K$.

Definition 2.6. Any linear map $D : C_p^\infty \to \mathbb{R}^n$ that satisfies the *Leibniz rule*

$$D_v(fg) = (D_v f)g(p) + f(p)(D_v g)$$

is called a *derivation at p* or a *point derivation* of C_p^∞. The set of all derivations at p is denoted $\mathcal{D}_p(\mathbb{R}^n)$.

Example 2.7. Directional derivatives are derivations. To see this, suppose $v \in T_p(\mathbb{R}^n)$ and f, g are C^∞ functions on a neighborhood of p. Then

$$\begin{aligned} D_v(fg) &= \sum_1^n v^i \frac{\partial(fg)}{\partial x^i}(p) \\ &= \left(\sum_1^n v^i \frac{\partial f}{\partial x^i}(p)\right) g(p) + f(p) \left(\sum_1^n v^i \frac{\partial g}{\partial x^i}(p)\right) \\ &= (D_v f)g(p) + f(p)(D_v g). \end{aligned}$$

Theorem 2.8. *The linear map $\phi : T_p(\mathbb{R}^n) \to \mathcal{D}_p(\mathbb{R}^n)$ defined by*

$$\phi(v) = D_v = \sum_1^n v^i \frac{\partial}{\partial x^i}\bigg|_p$$

is an isomorphism of vector spaces.

Thus if we make the identification $T_p(\mathbb{R}^n) \simeq \mathcal{D}_p(\mathbb{R}^n)$, the standard basis $\{e_1, \ldots, e_n\}$ of $\mathbb{R}^n$ becomes the standard basis $\{\partial/\partial x^1|_p, \ldots, \partial/\partial x^n|_p\}$ for $\mathcal{D}_p(\mathbb{R}^n)$. Similarly, we may identify the tangent vector $v = \langle v^1, \ldots, v^n \rangle$ in $T_p(\mathbb{R}^n)$ with the directional derivative

$$v = \sum_1^n v^i \frac{\partial}{\partial x^i}\bigg|_p .$$

2.4 Vector Fields

Definition 2.9. A *vector field* X on an open set U of $\mathbb{R}^n$ is a function that assigns to each $p \in \mathbb{R}^n$ a tangent vector X_p in $T_p(\mathbb{R}^n)$. Since $\{\partial/\partial_x^i|_p\}_{i=1}^n$ serves as a basis for $\mathcal{D}_p(\mathbb{R}^n)$ may write the tangent vector X_p as

$$X_p = \sum_1^n a^i(p) \frac{\partial}{\partial x^i}\bigg|_p .$$

We say that the vector field X is C^∞ or *smooth* on U if the coefficient functions a^i are all C^∞ on U.

It is easy to check that the set of all C^∞ functions on an open set U is an algebra over $\mathbb{R}$ (with the natural multiplication operation $(fg)(p) = f(p)g(p)$), and we will denote this algebra by $C^\infty(U)$.

Definition 2.10. Let R be a commutative ring with identity. An *R-module* is a set A with two binary operations, addition and scalar multiplication, such that

1. A is an abelian group under addition,
2. For $s, t \in R$ and $a, b \in A$
   - (a) (closure) $ta \in A$.
   - (b) (identity) If 1 is the multiplicative identity in R then $1a = a$.
   - (c) (associativity) $(st)a = s(ta)$.
   - (d) (distributivity) $(s + t)a = sa + ta$ and $s(a + b) = sa + sb$.

Note that if R is a field than an R-module is just a vector space over R.

Lemma 2.11. *Let U be open in $\mathbb{R}^n$, and let $\mathfrak{X}(U)$ denote the set of all C^∞ vector fields over U. Then $\mathfrak{X}(U)$ is a module over $C^\infty(U)$.*

2.5 Vector Fields as Derivations

Let U be open in $\mathbb{R}^n$. For $X \in \mathfrak{X}(U)$ and $f \in C^\infty(U)$ define Xf by

$$(Xf)(p) = X_p f.$$

Since $X_p = \sum_1^n a^i(p) \frac{\partial}{\partial x^i}\big|_p$ we have

$$(Xf)(p) = X_p f = \sum_1^n a^i(p) \frac{\partial f}{\partial x^i}(p).$$

Thus Xf is the function that at any point p returns the directional derivative of f in the direction of the tangent vector that X returns at p.

Definition 2.12. Let A be an algebra over a field K. A K-linear map $D : A \to A$ is called a *derivation* iff for all $a, b \in A$ it satisfies the Leibniz rule

$$D(ab) = (Da)b + a(Db).$$

The set of all derivations of A is denoted $\operatorname{Der}(A)$.

The reader should ponder for a moment the difference between derivations at p and derivations of the algebra C_p^∞. The former is a map from C_p^∞ into $\mathbb{R}$ whereas the latter is a map from C_p^∞ to C_p^∞.

Lemma 2.13. Der(A) *is a vector space over* K.

Proposition 2.14. *Let* U *be open in* $\mathbb{R}^n$. *If* $X \in \mathfrak{X}(U)$ *and* $f \in C^\infty(U)$ *then the map* $f \mapsto Xf$ *is* $\mathbb{R}$*-linear, and* $X(fg)$ *satisfies the Leibniz rule*

$$X(fg) = (Xf)g + f(Xg).$$

Thus the map $\phi : \mathfrak{X}(U) \to \mathrm{Der}(C^\infty(U)) : X \mapsto (f \mapsto Xf)$ *is a derivation of the algebra* $C^\infty(U)$.

3 Tensors

3.1 Dual Spaces

Definition 3.1. Let V and W be vector spaces over a field $\mathbb{F}$. A function $f : V \to W$ is said to be *linear* if $f(u + v) = f(u) + f(v)$ for all $u, v \in V$ and if $f(\lambda v) = \lambda f(v)$ for all $\lambda \in \mathbb{F}$. We denote by Hom(V, W) the vector space of all linear functions $f : V \to W$. We define the *dual space* of V by $V^\star = \mathrm{Hom}(V, \mathbb{F})$. △

Definition 3.2. Let V be a vector space with basis $e_1, \dots, e_n$. For any $i = 1, \dots, n$, define the function α^i by $\alpha^i(v) = v^i$, where $v = \sum_1^n v^i e_i$. This is well-defined because each $v \in V$ has a unique representation as a linear combination of vectors in the basis $e_1, \dots, e_n$. We call $\alpha^1, \dots, \alpha^n$ the *coordinate functions on* V *relative to the basis* $e_1, \dots, e_n$. △

Proposition 3.3. *Let* V *be a vector space with basis* $e_1, \dots, e_n$. *The coordinate functions* $\alpha^1, \dots, \alpha^n$ *relative to* $e_1, \dots, e_n$ *form a basis for the dual space* $V^\star$.

Proof. Let f_0 denote the zero element of $V^\star$, i.e. $f_0(v) = 0$ for all $v \in V$. Suppose there are constants $c_1, \dots, c_n$ such that $f_0 = \sum_1^n c_i \alpha^i$. Then for any $j = 1, \dots, n$ we have

$$0 = f_0(e_j) = (\sum_1^n c_i \alpha^i)(e_j) = \sum_1^n c_i \alpha^i(e_j) = \sum_1^n \delta_j^i = c_j.$$

Hence the coordinate functions $\alpha^1, \dots, \alpha^n$ are linearly independent. Now let $f \in V^\star$. For any $v \in V$, write $v = \sum_1^n v^i e_i$. By linearity of f, we have

$$f(v) = f(\sum_1^n v^i e_i) = \sum_1^n v^i f(e_i) = \sum_1^n f(e_i)\alpha^i(v).$$

Since this is true for arbitrary $v \in V$, we have $f = \sum_1^n f(e_i)\alpha^i$. That is, any $f \in V^\star$ can be expressed as a combination of the coordinate functions $\alpha^1, \dots, \alpha^n$, so the coordinate functions span $V^\star$. ☐

Definition 3.4. With notation as in Proposition 3.3, we say that $\alpha^1, \dots, \alpha^n$ is the basis for $V^\star$ that is *dual* to the basis $e_1, \dots, e_n$ of V. (Informally, the dual basis consists of the coordinate functions.)

Corollary 3.5. *If V is finite dimensional, then V and $V^\star$ have the same dimension.*

Theorem 3.6. *A collection of linear functions $f^1, \dots, f^n : V \to \mathbb{R}$ is the basis of $V^\star$ dual to the basis $e_1, \dots, e_n$ of V if and only if $f^i(e_j) = \delta^i_j$. (That is, the dual basis is characterized by the Kronecker delta property.)*

Proof.

($\Rightarrow$) Note that e_j has the unique representation $e_j = \sum_i v^i e_i$ where $v^i = 0$ for $i \neq j$ and $v^j = 1$. Since $f^1, \dots, f^n$ is the basis of $V^\star$ dual to the basis $e_1, \dots, e_n$ in V, then f^i are the coordinate functions relative to $e_1, \dots, e_n$. Hence $f^i(e_j) = \delta^i_j$.

($\Leftarrow$) Let $v \in V$ and write $v = \sum_i v^i e_i$. Then for $i = 1, \dots, n$, we have

$$f^i(v) = f^i(\sum_j v^j e_j) = \sum_j v^j f^i(e_j) = \sum_j v^j \delta^i_j = v^i.$$

Thus the functions $f^1, \dots, f^n$ are the coordinate functions relative to $e_1, \dots, e_n$. Hence they form the basis of $V^\star$ dual to the basis $e_1, \dots, e_n$ of V. ☐

Proposition 3.7. *Let V, W be vector spaces, and let $e_1, \dots, e_n$ be a basis for V. Suppose $f, g : V \to W$ and $f(e_i) = g(e_i)$ for all e_i. Then $f = g$.*

Proof. Let $v \in V$ and write $v = \sum_i v^i e_i$. Using linearity of f and g,

$$f(v) = f(\sum_i v^i e_i) = \sum_i v^i f(e_i) = \sum_i v^i g(e_i) = g(\sum_i v^i e_i) = g(v).$$

☐

Proposition 3.7 says that a linear function on a vector space is completely determined by its action on a basis.

3.2 Permutations

Definition 3.8. A *permutation* on the set $A = \{1, 2, \dots, k\}$ is a bijection $\sigma : A \to A$. The product $\tau\sigma$ of two permutations τ and σ is the composition $\tau \circ \sigma : A \to A$. The *cyclic permutation* $(a_1\, a_2 \cdots a_r)$ is the permutation σ such that $\sigma(a_1) = a_2$, $\sigma(a_2) = a_3, \dots, \sigma(a_{r-1}) = a_r$, $\sigma(a_r) = a_1$, and such that σ leaves all othere elements of A unchanged. The permutation $(a_1\, a_2 \cdots a_r)$ is called a *cycle of length* r or an r*-cycle*. A *transposition* is a cycle of the form $(a\, b)$ that swaps a and b, leaving all other elements of A fixed. An *inversion* in a permutation σ is an ordered pair $(\sigma(i), \sigma(j))$ such that $i < j$ but $\sigma(i) > \sigma(j)$.

Proposition 3.9. *Let σ be a permutation. Then σ can be written as a product of disjoint cycles. Furthermore, σ can be written as a product of transpositions, and any two such representations of σ as a product of transpositions either both consist of an odd number of transpositions, or both consist of an even number of transpositions.*

Definition 3.10. The *sign* of a permutation σ, denoted $\text{sgn}(\sigma)$, is +1 or -1 depending as to whether σ can be written as an even or odd number of transpositions. This is well-defined by Proposition 3.9.

Theorem 3.11.

1. *The set of all permutations on a set forms a group with the operation of function composition. We denote by S_k the group of all permutations on $\{1, 2, \dots, k\}$.*

2. *A permutation is even (odd) if and only if it has an even (odd) number of inversions. In fact, $sgn(\sigma) = (-1)^{\#\text{ of inversions in }\sigma}$.*

3. *For any two permutations σ and τ, we have $sgn(\sigma\tau) = sgn(\sigma)sgn(\tau)$.*

3.3 Tensors as Multilinear Functions

Definition 3.12. Let V be a vector space over the field $\mathbb{F}$ and k a positive integer. A k*-tensor* (or k*-linear function*) on V is a map $f : V^k \to \mathbb{R}$ that is linear in each variable separately, i.e.

$$f(\dots, au + bv, \dots) = af(\dots, u, \dots) + bf(\dots, v, \dots)$$

for any scalars $a, b \in \mathbb{F}$. We call k the *degree* of f. If f has degree 1, 2, or 3, we may also refer to it as a *linear*, *bilinear*, or *trilinear* function, respectively. The vector space of all k-tensors on V is denoted $L_k(V)$. △

Remark 3.13. The domain of a k-tensor is V^k, but for simplicitly we just refer to a k-tensor on V. We could increase clarity (at the cost of redundancy) by referring to a k-tensor on V^k.

3.4 Permutations Acting on k-Tensors

Definition 3.14. Given a k-tensor f on V and a permutation $\sigma \in S_k$ we define the k-tensor σf on V by

$$(\sigma f)(v_1, \dots, v_k) = f(v_{\sigma(1)}, \dots, v_{\sigma(k)}).$$

A k-tensor f is called *symmetric* if $\sigma f = f$ for all $\sigma \in S_k$, and it is called *alternating* if $\sigma f = (\operatorname{sgn} \sigma) f$, where $\operatorname{sgn} \sigma$ is the sign of the permutation σ. Let $S_k(V)$ denote the set of all symmetric k-tensors on V and $A_k(V)$ denote the set of all alternating k-tensors on V. We also refer to the elements of $A_k(V)$ as *k-covectors* or *multicovectors* on V. For $k = 0$ we define a 0-covector to be a constant, so that $A_0(V) = \mathbb{R}$. A 1-covector is simply called a *covector*. △

Note: The only permutation in S_1 is the identity permutation, which has sign $+1$. Hence all linear functions (i.e. 1-tensors) are alternating, and $L_1(V) = A_1(V) = V^\star$.

Example 3.15.

1. Let $V = \mathbb{R}^3$ and $k = 2$. The ordinary Euclidean inner product function $\langle v, w \rangle = \sum_1^3 v^i w^i$ is a symmetric 2-tensor on V. Indeed,

   $$\langle au + bv, w \rangle = \sum_1^3 (au^i + bv^i) w^i = a \sum_1^3 u^i w^i + b \sum_1^3 v^i w^i = a \langle u, w \rangle + b \langle v, w \rangle .$$

   Similarly, we can see that $\langle u, av + bw \rangle = a \langle u, v \rangle + b \langle u, w \rangle$. Thus $\langle \cdot, \cdot \rangle$ is a 2-tensor on V. Now let σ be the transposition $(1\ 2)$ in S_2. Then $\sigma \langle \cdot, \cdot \rangle (v, w) = \langle w, v \rangle$. Thus it is symmetric.

2. Let V be a vector space of dimension n over $\mathbb{R}$. The determinant function $\det : V^n \to \mathbb{R}$ is an alternating n-tensor on V^n. △

Definition 3.16. If G is a group and X is a set, a map

$$G \times X \to X : (\sigma, x) \mapsto \sigma \cdot x$$

is called a *left group action* (or just *left action*) of G on X if for any $\sigma, \tau \in G$ and $x \in X$ we have

1. $1 \cdot x = x$ where 1 is the identity in G, and
2. $\tau \cdot (\sigma \cdot x) = (\tau\sigma) \cdot x$.

A *right group action* (or just *right action*) is defined similarly.

Proposition 3.17. *The map $S_k \times L_k(V) \to L_k(V) : (\sigma, f) \mapsto \sigma f$ is a left group action on V. Moreover, it is linear in f. That is, $\sigma(af + bg) = a\sigma f + b\sigma g$.*

Proof. Let $e \in S_k$ be the identity permtuation. Then $e(i) = i$ for all indices i, so

$$(ef)(v_1, \dots, v_k) = f(v_{e(1)}, \dots, v_{e(k)}) = f(v_1, \dots, v_k).$$

Hence $ef = f$. Now let $\sigma, \tau \in S_k$. Then

$$\begin{aligned} \sigma(\tau f) &= (\tau f)(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \\ &= (\tau f)(w_1, \dots, w_k) \qquad (w_i := v_{\sigma(i)}) \\ &= f(w_{\tau(1)}, \dots, w_{\tau(k)}) \\ &= f(v_{\sigma(\tau(1))}, \dots, v_{\sigma(\tau(k))}) \\ &= (\sigma\tau)f. \end{aligned}$$

To establish the linearity, let a, b be scalars and $f, g \in L_k(V)$. Then

$$\begin{aligned} (\sigma(af + bg))(v_1, \dots, v_k) &= (af + bg)(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \\ &= af(v_{\sigma(1)}, \dots, v_{\sigma(k)}) + bg(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \\ &= a(\sigma f)(v_1, \dots, v_k) + b(\sigma g)(v_1, \dots, v_k). \end{aligned}$$

Hence $\sigma(af + bg) = a\sigma f + b\sigma g$. □

3.5 The Tensor Product

Definition 3.18. Given a k-tensor f and an ℓ-tensor g on V, define their *tensor product* to be the $k+\ell$-tensor $f \otimes g$ given by

$$(f \otimes g)(v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+\ell}) = f(v_1, \ldots, v_k)g(v_{k+1}, \ldots, v_{k+\ell}).$$

△

Example 3.19. Let $V = \mathbb{R}^n$. Define $f \in L_2(V)$ by $f(u, v) = u + 3v$ and define $g \in L_3(V)$ by $g(u, v, w) = 2u - v + 4w$. Then

$$(f \otimes g)(v, w, x, y, z) = f(v, w)g(x, y, z) = (u + 3v)(2x - y + 4z).$$

△

Example 3.20. Let $V = \mathbb{R}^n$. Let α^i for $i = 1, 2, \ldots, n$ denote the coordinate functions. That is, for any $v = (v^1, \ldots, v^n) \in V$, we have $\alpha^i(v) = v^i$. We may express the Euclidean inner product as a tensor product as follows.

$$\langle u, v \rangle = \sum_1^n u^i v^i = \sum_1^n \alpha^i(v)\alpha^i(w) = \sum_1^n (\alpha^i \otimes \alpha^i)(v, w).$$

Thus

$$\langle \cdot, \cdot \rangle = \sum_1^n \alpha^i \otimes \alpha^i.$$

△

Proposition 3.21. *The tensor product is associative. That is, if f, g, h are tensors on V, then*

$$f \otimes (g \otimes h) = (f \otimes g) \otimes h.$$

□

3.6 Multi-Index Notation

Definition 3.22. Let A be an indexing set. A *length k multi-index* over A is an ordered list of length k with entries from A. Two multi-indices I and J are said to be equal if they are equal as ordered lists (i.e. they have the same length and the same entries in the same order). If $I = (i_1, \ldots, i_k)$, we say that I has *repeats* if $i_j = i_\ell$ for some $j \neq \ell$, and we say that I is *ascending* if $i_1 < \cdots < i_k$. (Thus an ascending multi-index has no repeats.) If $I = (i_1, \ldots, i_k)$ and $\sigma \in S_k$ then we define $\sigma I = (i_{\sigma(1)}, \ldots, i_{\sigma(k)})$. When the indexing set is obvious we do not specify it explicitly. △

Example 3.23. Suppose $e_1, \ldots, e_7$ are vectors in a vector space V. Then $I = (2, 3, 5)$, $J = (6, 4, 7)$, and $K = (1, 1, 1)$ are each length 3 multi-indices over the indexing set $A = \{1, \ldots, 7\}$. Note that I is ascending whereas J and K are not. In this context, the indexing set $A = \{1, \ldots, 7\}$ is obvious, and we would not normally specify it explicitly. If σ is the permutation $(2\ 1\ 3) \in S_3$ then $\sigma J = (4, 6, 7)$. △

Proposition 3.24. *Let $e_1, \ldots, e_n$ be a basis of the vector space V, and let $\alpha^1, \ldots, \alpha^n$ be the corresponding dual basis of $V^\star$. For a multi-index $I = (i_1, \ldots, i_k)$, temporarily define*[1] *the notation $\alpha^I = \alpha^{i_1} \otimes \cdots \otimes \alpha^{i_k}$ and $e_I = (e_{i_1}, \ldots, e_{i_k})$.*

1. *If I and J are two multi-indices of length k, then $\alpha^I(e_J) = \delta^I_J$.*
2. *Let $\mathcal{A}$ be the set of all tensor products α^I as I runs over all multi-indices of length k. Then $\mathcal{A}$ is a basis for $L_k(V)$.*

Proof.

1. Let $I = (i_1, \ldots, i_k)$ and $J = (j_1, \ldots, j_k)$. Then

   $$\alpha^I(e_J) = (\alpha^{i_1} \otimes \cdots \otimes \alpha^{i_k})(e_{j_1}, \ldots, e_{j_k}) = \alpha^{i_1}(e_{j_1}) \cdots \alpha^{i_k}(e_{j_k}) = \delta^{i_1}_{j_1} \cdots \delta^{i_k}_{j_k} = \delta^I_J,$$

   where the third equality followed by Theorem 3.6. The last equality follows because $\delta^{i_1}_{j_1} \cdots \delta^{i_k}_{j_k}$ equals 1 if and only if $i_t = j_t$ for all $t = 1, \ldots, k$ and equals zero otherwise.

2. Let M be the set of all multi-indices of length k over $\{1, \ldots, n\}$. Let f_0 be the zero element of $L_k(V)$. That is, $f_0(v_1, \ldots, v_k) = 0$ for any $v_1, \ldots, v_k$ in V. Now suppose $\sum_{I \in M} c_I \alpha^I = f_0$, for constants $c_I \in \mathbb{R}$ (thus the sum has n^k terms in it). For any $J \in M$ we have

   $$0 = f_0(e_J) = (\sum_{I \in M} c_I \alpha^I)(e_J) = \sum_{I \in M} c_I \alpha^I(e_J) = \sum_{I \in M} c_I \delta^I_J = c_I,$$

   where the next to last equality followed by part 1. Thus the set $\mathcal{A}$ is linearly independent. Now let $f \in L_k(V)$. By analogy with the proof

---

[1] We will redefine this notation later in a slightly different way.

of Proposition 3.3, we make the ansatz $f = \sum_{I \in M} f(e_I)\alpha^I$. For any $J \in M$ we have

$$(\sum_{I \in M} f(e_I)\alpha^I)(e_J) = \sum_{I \in M} f(e_I)\alpha^I(e_J) = \sum_{I \in M} f(e_I)\delta^I_J = f(e_J).$$

Thus the function $\sum_{I \in M} f(e_I)\alpha^I$ agrees with the function f on e_J for all $J \in M$. Since $\{e_J : J \in M\}$ is a basis for $V \times \cdots \times V$ (k factors), Proposition 3.7 implies that $f = \sum_{I \in M} f(e_I)\alpha^I$. Thus $\mathcal{A}$ spans $L_k(V)$.

□

Corollary 3.25. *If V has dimension n, then $L_k(V)$ has dimension n^k.*

3.7 The Symmetrizing and Alternating Operators

Given any k-tensor on a vector space V, there is a standard way of obtaining a symmetric or alternating k-tensor from it. We introduce the symmetrizing and alternating operators to achieve this.

Theorem 3.26. *For any $f \in L_k(V)$, define Sf and Af by*

$$Sf = \sum_{\sigma \in S_k} \sigma f,$$
$$Af = \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\, \sigma f.$$

Then $Sf \in S_k(V)$ and $Af \in A_k(V)$. We call the map $S : L_k(V) \to S_k(V)$ the symmetrizing operator *and the map $A : L_k(V) \to A_k(V)$ the* alternating operator*.*

Proof. We will need the following lemma twice in our proof: For any fixed $\tau \in S_k$, the map $f : S_k \to S_k : \sigma \mapsto \tau\sigma$ is a bijection. This follows from the group properties of S_k. To wit, let $\sigma \in S_k$. Then $f(\tau^{-1}\sigma) = \tau(\tau^{-1}\sigma) = \sigma$, so f is surjective. If $f(\sigma_i) = f(\sigma_j)$ then $\tau\sigma_i = \tau\sigma_j$, so multiplying through by τ^{-1} we obtain $\sigma_i = \sigma_j$. Thus f is injective, and hence bijective.

Now let $f \in L_k(V)$ and consider $Sf = \sum_{\sigma \in S_k} \sigma f$. We must show that $\tau(Sf) = Sf$ for any $\tau \in S_k$. But

$$\tau(Sf) = \tau \sum_{\sigma \in S_k} \sigma f = \sum_{\sigma \in S_k} \tau(\sigma f) = \sum_{\sigma \in S_k} (\tau\sigma) f = \sum_{\nu \in S_k} \nu f = Sf,$$

where the third equality followed by Proposition 3.17, and the fourth proposition followed from the lemma. Hence Sf is indeed symmetric.

Finally, consider $Af = \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\, \sigma f$. We need to show that $\tau(Af) = (\operatorname{sgn} \tau)(Af)$ for any $\tau \in S_k$. But

$$\begin{aligned}\tau(Af) = \tau \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\, \sigma f &= \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\, \tau\sigma f = (\operatorname{sgn} \tau) \sum_{\sigma \in S_k} (\operatorname{sgn} \tau)\, (\operatorname{sgn} \sigma)\, \tau\sigma f \\ &= (\operatorname{sgn} \tau) \sum_{\sigma \in S_k} (\operatorname{sgn} (\tau\sigma)) \tau\sigma f \ = (\operatorname{sgn} \tau) \sum_{\nu \in S_k} (\operatorname{sgn} \nu)\, \nu f = (\operatorname{sgn} \tau)\, (Af),\end{aligned}$$

where in going to the third equality I multipled by $(\operatorname{sgn} \tau)\, (\operatorname{sgn} \tau) = 1$. In going to the fourth equality I used the fact that $(\operatorname{sgn} \tau)\, (\operatorname{sgn} \sigma) = (\operatorname{sgn} \tau\sigma)$, in going to the fifth equality I used the lemma. Hence Af is indeed alternating. $\square$

3.8 The Wedge Product

Given an alternating k-tensor and an alternating ℓ-tensor g, we would like to form their product in a way that results in an alternating $k + \ell$-tensor. This is accomplished by the wedge product.

Definition 3.27. Let $f \in A_k(V)$, $g \in A_\ell(V)$. Define their *wedge product* as follows.

$$f \wedge g = \frac{1}{k!\ell!} A(f \otimes g).$$

Written out explicitly, this becomes

$$\begin{aligned}&(f \wedge g)(v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+\ell}) \\ &\qquad = \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} (\operatorname{sgn} \sigma) f(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) g(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+\ell)}).\end{aligned}$$

$\triangle$

Remark 3.28. At first glance, the restriction that f and g be alternating seems unnecessary, since $\frac{1}{k!\ell!} A(f \otimes g)$ is alternating whether or not f and g are. However, without the stipulation that f and g are alternating, some of the nice properties of the wedge product that we need later would no longer hold true.

The $\frac{1}{k!\ell!}$ factor preceding the sum accounts for repeated values. For example, let $\sigma_1, \ldots, \sigma_{k!}$ be the $k!$ permutations in $S_{k+\ell}$ that permute the indices $1, \dots, k$ while leaving the indices $k+1, \dots, k+\ell$ fixed. If we view these special permutations as members of S_k and note that $(\operatorname{sgn}\sigma)\sigma f = (\operatorname{sgn}\sigma)(\operatorname{sgn}\sigma)f = f$ (since f is alternating), it follows that $\sigma_1 f, \dots, \sigma_{k!} f$ each contribute the same value to the sum when evaluated on $(v_1, \dots, v_k)$. Thus we divide the sum by $k!$ in the definition. Similarly with the $\frac{1}{\ell!}$ factor.

Finally, note that calculating the wedge product directly from the definition, even for small examples, is cumbersome to say the least. For example, if f and g are both 2-tensors on a vector space V, then computing their wedge product involves summing over permutations from $S_{2+2} = S_4$. Since $|S_4| = 4! = 24$, there are 24 terms in the sum! However, we can greatly simplify matters by introducing (k, ℓ)-shuffles.

Definition 3.29. A permutation $\sigma \in S_{k+\ell}$ is called a (k, ℓ)-shuffle if

$$\sigma(1) < \cdots < \sigma(k) \quad \text{and} \quad \sigma(k+1) < \cdots < \sigma(k+\ell).$$

We may similarly define $(k_1, k_2, \dots, k_n)$-shuffles in $S_{k_1+k_2+\cdots+k_n}$. △

Remark 3.30. Note that no stipulation is made about the relationship between $\sigma(k)$ and $\sigma(k+1)$. (Indeed, if we required $\sigma(k) < \sigma(k+1)$ then the only (k, ℓ)-shuffle would be the identity permutation. Think about it.) Also note that the set of all $(1, 1)$-shuffles in S_2 is just S_2 itself, the set of all $(1, 1, 1)$-shuffles in S_3 is just S_3 itself, etc. Part 4 of the Proposition 3.31 tells us how to use (k, ℓ)-shuffles to evaluate the wedge product.

Proposition 3.31 (Properties of the wedge product)**.** *The wedge product has the following properties:*

1. *It is anti-commutative in the sense that if $f \in A_k(V)$, $g \in A_\ell(V)$, then $f \wedge g = (-1)^{k\ell} g \wedge f$.*
2. *If $f \in A_k(V)$ and k is odd, then $f \wedge f = 0$.*
3. *It is associative, i.e. $f \wedge (g \wedge h) = (f \wedge g) \wedge h$.*
4. *If $f \in A_k(V)$ and $g \in A_\ell(V)$, then*

$$\begin{aligned}(f \wedge g)&(v_1, \dots, v_k, v_{k+1}, \dots, v_{k+\ell}) \\ &= \sum_{\substack{(k,\ell)\text{-shuffles} \\ \sigma}} (\mathit{sgn}\,\sigma) f(v_{\sigma(1)}, \dots, v_{\sigma(k)}) g(v_{\sigma(k+1)}, \dots, v_{\sigma(k+\ell)})\end{aligned}$$

5. *Let $f^i \in A_{k_i}(V)$ for $i = 1, \dots, r$. Then*

$$f^1 \wedge \cdots \wedge f^r = \frac{1}{k_1! \cdots k_r!} A(f^1 \otimes \cdots \otimes f^r).$$

6. *If $f^1, \dots, f^k \in L_1(V)$ and $v_1, \dots, v_k \in V$, then*

$$(f^1 \wedge \cdots \wedge f^k)(v_1, \dots, v_k) = \det[f^i(v_j)],$$

where $[f^i(v_j)]$ denotes the matrix whose (i, j)-entry is $f^i(v_j)$.

Definition 3.32. We call the matrix in part 6 of Proposition 3.31 the *wedge product matrix* associated with $f^1, \dots, f^k$ and $v_1, \dots, v_k$. We typically denote this matrix by $\Lambda(\alpha^1, \dots, \alpha^k, v_1, \dots, v_k)$, or just Λ. △

Example 3.33. Let $f \in A_2(V)$ and $g \in A_2(V)$. Using $(2, 2)$-shuffles in S_4 and property 4 of the preceding proposition, we have

$$\begin{aligned}(f \wedge g)(v_1,&v_2, v_3, v_4)\\ &= f(v_1, v_2)g(v_3, v_4) - f(v_1, v_3)g(v_2, v_4) - f(v_1, v_4)g(v_2, v_3)\\ &\quad + f(v_2, v_3)g(v_1, v_4) - f(v_2, v_4)g(v_1, v_3) + f(v_3, v_4)g(v_1, v_2).\end{aligned}$$

△

Remark 3.34. For wedge products of 1-tensors, the determinant method applies (part 6 of proposition 3.31), and it is even easier than using shuffles. The reason is because calculating determinants is a mechanical process that requires little mental thought, and it forces us to keep our work organized in neat rows and columns. On the other hand, calcualting wedge products either from the definition or by using shuffles requires us to form a list of permutations and determine their signs, something that's not easy to do mentally and is prone to sign errors.

Example 3.35. Let V be a vector space, $v_1, \dots, v_5$ be vectors in V, and $f^1, \dots, f^5$ be linear functions on V. Consider the multi-indices $I = (1, 2)$, $J = (2, 2)$, $K = (4, 1, 5)$, and $L = (1, 3, 2)$. Using part 6 of Proposition

3.31 we get

$$
\begin{aligned}
f^I(v_I) &= (f^1 \wedge f^2)(v_1, v_2) = \det \begin{bmatrix} f^1(v_1) & f^1(v_2) \\ f^2(v_1) & f^2(v_2) \end{bmatrix} \\
f^J(v_I) &= (f^2 \wedge f^2)(v_1, v_2) = \det \begin{bmatrix} f^2(v_1) & f^2(v_2) \\ f^2(v_1) & f^2(v_2) \end{bmatrix} \\
f^K(v_L) &= (f^4 \wedge f^1 \wedge f^5)(v_1, v_3, v_2) = \det \begin{bmatrix} f^4(v_1) & f^4(v_3) & f^4(v_2) \\ f^1(v_1) & f^1(v_3) & f^1(v_2) \\ f^5(v_1) & f^5(v_3) & f^5(v_2) \end{bmatrix}.
\end{aligned}
$$

To see the superiority of the determinant method for wedge products of 1-tensors, consider expanding $f^K(v_L)$ using $(1, 1, 1)$-shuffles on $S_{1+1+1} = S_3$ instead of determinants. You'll quickly see the determinant method is better for hand calculations. △

Definition 3.36. Let V be a vector space, $e_1, \dots, e_n$ be vectors in V, and $\alpha^1, \dots, \alpha^n$ be 1-tensors (i.e. linear functions) on V. Let $I = (i_1, \dots, i_k)$ and $J = (j_1, \dots, j_k)$ be length k multi-indices over $\{1, \dots, n\}$. Define the symbols α^I, e_J, and $\alpha^I(e_J)$ by

$$
\begin{aligned}
\alpha^I &= \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k} \\
e_J &= (e_{j_1}, \dots, e_{j_k}) \\
\alpha^I(e_J) &= (\alpha^{i_1} \wedge \ldots \wedge \alpha^{i_k})(e_{j_1}, \dots, e_{j_k}).
\end{aligned}
$$

△

Note: Recall that $L_1(V) = A_1(V) = V^\star$. Thus $\alpha^1, \dots, \alpha^k$ are alternating, and it makes sense to form their wedge product.

Proposition 3.37. *Let V be a vector space with basis $e_1, \dots, e_n$. Let $\alpha^1, \dots, \alpha^n$ be the basis for $V^\star$ dual to the basis $e_1, \dots, e_n$ of V. Suppose $I = (i_1, \dots, i_k)$ and $J = (j_1, \dots, j_k)$ are two length k multi-indices over $\{1, \dots, n\}$. Then*

1. *$\alpha^I(e_J) = 0$ if I has any repeats.*
2. *$\alpha^I(e_J) = 0$ if $k > n$.*

Assuming that I has no repeats and $k \leq n$, then

3. *$\alpha^I(e_I) = 1$.*

4. *We have the following Kronecker delta like property:*

$$\alpha^I(e_J) = \begin{cases} \operatorname{sgn}\sigma, & \text{if } I = \sigma J \text{ for some } \sigma \in S_k \\ 0, & \text{otherwise.} \end{cases}.$$

5. *If I and J are ascending multi-indices, then we have the full Kronecker delta property, namely*

$$\alpha^I(e_J) = \delta^I_J = \begin{cases} 1, & I = J \\ 0, & I \neq J \end{cases}.$$

Proof.

1. Since I has repeats, at least two rows of the associated wedge product matrix Λ are the same. Hence $\alpha^I(e_J) = \det\Lambda = 0$.

2. If $k > n$ then I must have repeats (pigeonhole principle). Apply part 1.

3. Since $\alpha^1, \ldots, \alpha^n$ is the basis for $V^\star$ dual to the basis $e_1, \ldots, e_n$ of V, then Theorem 3.6 implies $\alpha^{i_r}(e_{j_s}) = \delta^{i_r}_{j_s}$ for any $1 \leq r, s \leq k$. Thus the wedge product matrix Λ associated to $\alpha^I(e_I)$ is the $k \times k$ identity matrix. Hence $\alpha^I(e_I) = \det\Lambda = 1$.

4. If $I = \sigma J$ for some $\sigma \in S_k$, then since $\alpha^I = \alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}$ is an alternating k-tensor, Definition 3.14 gives

$$\operatorname{sgn}\sigma \cdot [\alpha^I(e_J)] = (\sigma\alpha^I)(e_J) = \alpha^I(e_{\sigma J}) = \alpha^I(e_I) = 1,$$

   where we used part 3 in the last equality. Thus $\alpha^I(e_J) = \frac{1}{\operatorname{sgn}\sigma} = \operatorname{sgn}\sigma$ (since $\operatorname{sgn}\sigma = \pm 1$). On the other hand, if $I \neq \sigma J$ for any $\sigma \in S_k$, then since I has no repeats it follows that there is an index i_w in $I \setminus J$. Then $\alpha^{i_w}(e_{j_s}) = \delta^{i_w}_{j_s} = 0$ for all $1 \leq s \leq k$. So the wth row of the associated wedge product matrix Λ is a zero row, whence $\alpha^I(e_J) = \det\Lambda = 0$.

5. If $I = J$ then apply part 3 of this proposition. If $I \neq J$, then since I and J are both ascending it follows there is no $\sigma \in S_k$ such that $I = \sigma J$. Apply part 4.

$\square$

Example 3.38. To see parts 3 and 4 of Proposition 3.37 in action, let $\alpha^1, \dots, \alpha^5$ be the basis of $V^\star$ dual to the basis $e_1, \dots, e_5$ of the vector space V. Let $I = (2,3,1)$, $J = (3,2,1)$, and $K = (2,3,4)$. Notice that $I = \sigma J$ where σ is the permutation $(2\ 1\ 3)$ (with sign -1) in S_3, and that $I \neq \sigma K$ for any $\sigma \in S_3$. Thus by Proposition 3.37, we expect that $\alpha^I(e_I) = 1$, $\alpha^I(e_J) = \operatorname{sgn}\sigma = -1$, and $\alpha^I(e_K) = 0$. Indeed, we have

$$\alpha^I(e_I) = \det \begin{bmatrix} \alpha^2(e_2) & \alpha^2(e_3) & \alpha^2(e_1) \\ \alpha^3(e_2) & \alpha^3(e_3) & \alpha^3(e_1) \\ \alpha^1(e_2) & \alpha^1(e_3) & \alpha^1(e_1) \end{bmatrix} = \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 1,$$

$$\alpha^I(e_J) = \det \begin{bmatrix} \alpha^2(e_3) & \alpha^2(e_2) & \alpha^2(e_1) \\ \alpha^3(e_3) & \alpha^3(e_2) & \alpha^3(e_1) \\ \alpha^1(e_3) & \alpha^1(e_2) & \alpha^1(e_1) \end{bmatrix} = \det \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = -1,$$

and

$$\alpha^I(e_K) = \det \begin{bmatrix} \alpha^2(e_2) & \alpha^2(e_3) & \alpha^2(e_4) \\ \alpha^3(e_2) & \alpha^3(e_3) & \alpha^3(e_4) \\ \alpha^1(e_2) & \alpha^1(e_3) & \alpha^1(e_4) \end{bmatrix} = \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0.$$

3.9 A Basis for k-Covectors

Proposition 3.39. *Let* $e_1, \dots, e_n$ *be a basis for the vector space* V*, and let* $\alpha^1, \dots, \alpha^n$ *be the basis of* $V^\star$ *dual to the basis* $e_1, \dots, e_n$ *of* V*. Let* $\mathcal{I}_k$ *denote the set of all ascending length* k *multi-indices over* $\{1, \dots, n\}$*. Then the set* $\{\alpha^I : I \in \mathcal{M}_k\}$ *forms a basis for the vector space* $A_k(V)$ *of all alternating* k*-tensors on* V*.*

Corollary 3.40. *If* V *has dimension* n*, the dimension of* $A_k(V)$ *is* $\binom{n}{k}$*.*

The reader should take a moment to contrast Proposition 3.39 and its corollary with Proposition 3.24 and its corollary.

4 Differential Forms on $\mathbb{R}^n$

4.1 Differential 1-Forms and the Differential of a Function

4.2 Differential k-Forms

4.3 Differential Forms as Multilinear Functions on Vector Fields

4.4 The Exterior Derivative

4.5 Closed Forms and Exact Forms

4.6 Applications to Vector Calculus

$$\begin{tikzcd}
\Omega^0(U) \arrow[r, "d"] \arrow[d, "\simeq"'] & \Omega^1(U) \arrow[r, "d"] \arrow[d, "\simeq"'] & \Omega^2(U) \arrow[r, "d"] \arrow[d, "\simeq"'] & \Omega^3(U) \arrow[d, "\simeq"'] \\
C^\infty(U) \arrow[r, "\text{grad}"'] & \mathfrak{X}(U) \arrow[r, "\text{curl}"'] & \mathfrak{X}(U) \arrow[r, "\text{div}"'] & C^\infty(U)
\end{tikzcd}$$

4.7 Convention on Subscripts and Superscripts

5 Manifolds

5.1 Topological Manifolds

Definition 5.1. A topological space $(M, \mathfrak{T})$ is said to be *locally Euclidean of dimension* n if to each $p \in M$ there is a neighborhood U of p that is homemorphic to an open subset of $\mathbb{R}^n$. Where the underlying topology $\mathfrak{T}$ is understood or not important, we will denote the space simply by M.

Theorem 5.2. (Brouwer's Theorem on Invariance of Domain) *If U is an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^n$ is an injective continuous map, then $V = f(U)$ is open and f is a homeomorphism between U and V.*

It follows from Brouwer's theorem that an open set U in $\mathbb{R}^n$ cannot be homeomorphic to an open set V in $\mathbb{R}^m$ whenever $n \neq m$ (verify!). Thus in the definition of locally Euclidean, the integer n is uniquely determined by the topological space under consideration.

Definition 5.3. Given a topological space M, a *coordinate chart* (or just *chart*) on M is a pair (U, ϕ), where U is open in M and ϕ is a homeomorphism from U onto an open subset of $\mathbb{R}^n$. In this context, we call U a *coordinate neighborhood* and ϕ a *coordinate map* on U, and for a point $p \in M$ we say that the chart (U, ϕ) contains p if $p \in U$.

Definition 5.4. A *topological n-manifold* is a topological space M such that

1. M is Hausdorff.
2. M is second countable.
3. M is locally Euclidean of dimension n.

More briefly, a topological n-manifold is a second countable Hausdorff space that is locally Euclidean of dimension n.

The reader should note that some authors do not include the second countability condition on a topological manifold. When second countability is not assumed, some of the results we derive will no longer hold. For example, a locally Euclidean Hausdorff space need not be paracompact, in contrast to part 5 of Proposition 5.6 below.

Example 5.5. The space $\mathbb{R}^n$ with its usual (metric) topology is a topological manifold. Properties 1 and 2 of the definition are clearly satisfied with the metric topology. To see that $\mathbb{R}^n$ is locally Euclidean, let $p \in \mathbb{R}^n$ and let U be the open ball of radius 1 about p. Let $\phi : U \to U$ be the identity map on U. Then (U, ϕ) is ϕ is a homeomorphism from U to an open subset of $\mathbb{R}^n$, namely U itself! Thus $\mathbb{R}^n$ is a topological manifold.

Proposition 5.6. (Properties of topological n-manifolds.) *Let M be a topological n-manifold. Then:*

1. *M is locally path connected, hence locally connected.*
2. *The components of M are the same as its path components. Thus we may just speak of the components of M, and M is connected if and only if it is path connected.*

3. *The components of M are clopen (i.e. open and closed) and are countable in number.*

4. *M is locally compact.*

5. *M is paracompact.*

6. *M is normal.*

7. *M is metrizable.*

8. *If M is compact, then M can be embedded in $\mathbb{R}^N$ for some integer N.*

Proof.

1. We must show that for every $p \in M$ and neighborhood U of p there exists a path connected neighborhood U' of p with $U' \subset U$. Let $p \in M$. Then there is a neighborhood U of p, an open ball V in $\mathbb{R}^n$, and a homeomorphism ϕ from U onto V. The open ball $V = \phi(U)$ is path connected in $\mathbb{R}^n$. Since ϕ is a homeomorphism, ϕ^{-1} is continuous. Hence $U = \phi^{-1}(\phi(U)) = \phi^{-1}(V)$ is path connected in M since the continuous image of a path connected set is path connected. With $U' = U$, we see that M is locally path connected (and hence locally connected).

2. See Theorem 25.5 in Munkres [3].

3. Let $\mathcal{C}$ be a component of M and $p \in \mathcal{C}$. Since M is locally connected by part 1 of this theorem, there is a connected neighborhood U containing p. Since U is connected, either $U \subset \mathcal{C}$ or $U \cap \mathcal{C} = \varnothing$. The latter is impossible because both U and $\mathcal{C}$ contain p. Hence $U \subset \mathcal{C}$ and so $\mathcal{C}$ is open. Next, since $\mathcal{C}$ is connected it follows by Theorem 23.4 in Munkres [3] that $\overline{\mathcal{C}}$ is connected as well. By Theorem 25.1 in Munkres [3], it follows that $\overline{\mathcal{C}} \subset \mathcal{C}$. Hence $\mathcal{C}$ is closed.

4. See Corollary 1.7 of Lee [2].

5. Actually, all second countable locally compact Hausdorff spaces are paracompact. This is a good exercise.

6. Since M is both Hausdorff and paracompact, it is normal by Theorem 41.1 in Munkres [3].

7. Since M is normal it is regular. It is also second countable. By Urysohn's metrization theorem (Theorem 34.1 in Munkres [3]), M is metrizable.

8. The proof is somewhat involved and uses a partition of unity argument. See Theorem 36.2 in Munkres [3].

□

Sometimes we want to show that a given topological space is *not* a topological manifold. The following proposition can be useful in this regard.

Proposition 5.7. *A topological space M is locally Euclidean of dimension n if and only if to every point $p \in M$ and every open ball B in $\mathbb{R}^n$ there is a neighborhood U of p and a homeomorphism ϕ from U onto B.*

Proof. Let $p \in M$. Since M is locally Euclidean, there exists a neighborhood U of p and a homeomorphism ϕ from U onto an open subset V of $\mathbb{R}^n$, with $\phi(p) \in V$. Since V is open, there is a ball B' centered at $\phi(p)$ such that $B' \subset V$. Since $\phi : U \to V$ is a homeomorphism, ϕ^{-1} is continuous and $\phi^{-1}(B') \subset U$. Thus if we let ψ be the restriction of ϕ to $\phi^{-1}(B')$, i.e. $\psi = \phi^{-1}|_{B'}$, then ψ serves as a homeomorphism from the neighborhood $\phi^{-1}(B')$ of p onto the open ball B' of $\mathbb{R}^n$. Since translations and dilations of $\mathbb{R}^n$ are homeomorphisms, it follows that we may compose ψ with translations and dilations so that B' is homeomorphic to B. The converse is obvious. □

In using the preceding proposition, it is often useful to let B be the unit ball centered at $\phi(p)$ or centered at 0.

Example 5.8. Let M be the cross depicted in Figure 5.1, equipped with the subspace topology on $\mathbb{R}^2$. Let p denote the center of the cross.

If M is locally Euclidean of dimension n at p, then there exists a neighborhood U of p and a homeomorphism ϕ from U onto the open ball in $\mathbb{R}^n$ of radius 1 centered at $\phi(p)$. Now, ϕ restricts to a homeomorphism from $M \setminus \{p\}$ onto $B \setminus \{\phi(p)\}$. But $M \setminus \{p\}$ has four connected components, whereas $B \setminus \{\phi(p)\}$ has either one or two according as to whether $n \geq 2$ or $n = 1$, respectively. This is a contradiction. Thus M cannot be locally Euclidean at p, and so it is *not* a topological manifold.

In the same way, one can see that a sphere with hair is not a manifold. See Figure 5.2.

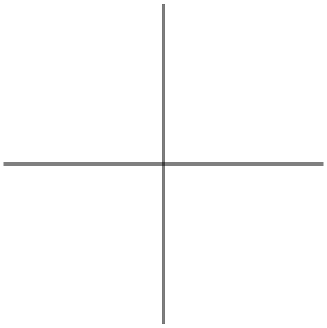


Figure 5.1: The cross is *not* a manifold.

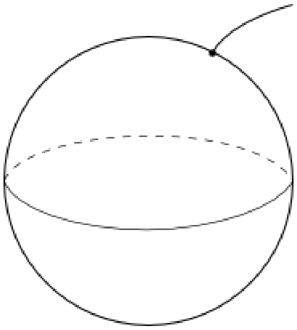


Figure 5.2: A sphere with hair is *not* a manifold.

Theorem 5.9. *An open subset of a topological manifold is itself a topologial manifold when given the subspace topology.*

Proof. Let M be a topological manifold and $U \subset M$ be open. Equip U with the subspace topology. The property of being Hausdorff and second countable are hereditary, so we need only check that U is locally Euclidean. Since M is locally Euclidean, for any $p \in U$ there is a coordinate chart (V, ϕ) with $p \in V$. The set $U \cap V$ is an open set in the subspace topology containing p, and if we let $\phi|_{U \cap V}$ denote the restriction of ϕ to $U \cap V$, then $\phi|_{U \cap V}$ is clearly a homeomorphism from $U \cap V$ into the open set $\phi(U \cap V)$ of $\mathbb{R}^n$. Since $p \in U$ was arbitrary, the result follows. □

The preceding theorem shows how abundant manifolds are. Every open subset of $\mathbb{R}^n$, for example, is a manifold when given the subspace topology.

5.2 Differentiable Manifolds

Definition 5.10. Two coordinate charts (U, ϕ) and (V, ψ) on a manifold are said to be C^∞-*compatible* (or just *compatible*) if $U \cap V$ nonempty implies $\phi \circ \psi^{-1} : \psi(U \cap V) \to \phi(U \cap V)$ and $\psi : \phi^{-1} : \phi(U \cap V) \to \psi(U \cap V)$ are C^∞. The maps $\phi \circ \psi^{-1}$ and $\psi \circ \phi^{-1}$ are called the *transition functions*.

The definition is illustrated in Figure 5.3.

Figure 5.3: Compatible charts and transition functions.

Consider two compatible charts (U, ϕ) and (V, ψ). Since U and V are open, so are $U \cap V$, $\phi(U \cap V)$, and $\psi(U \cap V)$. The transition functions $\phi \circ \psi^{-1}$ and $\psi \circ \phi^{-1}$ are homeomorphisms on $U \cap V$, being the composition of two homeomorphisms there. Moreover, they are C^∞ and inverses of each other, hence they are actually diffeomorphisms between $\psi(U \cap V)$ and $\phi(U \cap V)$.

Definition 5.11. Let M be a topological manifold. An *atlas* (or *differentiable (or C^∞ or smooth) structure*) on M is a family $\mathcal{U} = \{(U_\alpha, \varphi_\alpha)\}$ of coordinate charts such that:

1. The U_α cover M, i.e. $M = \bigcup U_\alpha$.
2. The charts (U_α, ϕ_α) and (U_β, ϕ_β) are compatible for any α, β.

An atlas $\mathcal{U}$ with the following additional property is called a *maximal atlas*:

3. If (V, ψ) is a chart compatible with every (U_α, ϕ_α) in $\mathcal{U}$, then (V, ψ) belongs to $\mathcal{U}$.

If (V, ψ) is a chart that is compatible with every chart in $\mathcal{U}$, then it is said to be *compatible with the atlas* $\mathcal{U}$.

Thus a maximal atlas is an atlas that contains all charts compatible with it. Equivalently, an atlas $\mathfrak{M}$ is maximal if whenever $\mathcal{U}$ is an atlas containing $\mathfrak{M}$ then $\mathcal{U} = \mathfrak{M}$.

Proposition 5.12. *Let $\mathcal{U} = \{(U_\alpha, \phi_\alpha)\}$ be an atlas on M. If two charts (V, ψ) and (W, σ) are compatible with $\mathcal{U}$, then they are compatible with each other.*

Proof. We show that $\sigma \circ \psi^{-1}$ is C^∞ on $\psi(V \cap W)$. That $\psi \circ \sigma^{-1}$ is C^∞ on $\sigma(V \cap W)$ is handled similarly. Let $p \in V \cap W$. (If $V \cap W = \varnothing$ the result is vacuously true.) Since $\mathfrak{U}$ is an atlas, it covers M. Thus there exists a chart (U_α, ϕ_α) in $\mathfrak{U}$ with $p \in V \cap W \cap U_\alpha$. Since (V, ψ) and (W, σ) are both compatible with (U_α, ϕ_α), it follows that (i) $\phi \circ \psi^{-1} : \psi(V \cap W \cap U_\alpha) \to \phi(V \cap W \cap U_\alpha)$ is C^∞, and (ii) $\sigma \circ \phi^{-1} : \phi(V \cap W \cap U_\alpha) \to \sigma(V \cap W \cap U_\alpha)$ is C^∞. Since $p \in V \cap W \cap U_\alpha$, we have that

$$\sigma \circ \psi^{-1} = (\sigma \circ \phi^{-1}) \circ (\phi \circ \psi^{-1})$$

is C^∞ at $\psi(p)$. Since the choice of $p \in V \cap W$ was arbitrary, we win. □

Definition 5.13. A *smooth* (or *differentiable* or C^∞-) *manifold* is a topological manifold equipped with a maximal atlas.

Theorem 5.14. *Let M be a topological manifold equipped with an atlas $\mathfrak{U} = \{(U_\alpha, \phi_\alpha)\}$. Then there is a unique maximal atlas $\mathfrak{M}$ containing $\mathfrak{U}$.*

Proof. Set $\mathfrak{M} = \mathfrak{U}$. Adjoin to $\mathfrak{M}$ any chart (V, ψ) on M that is compatible with $\mathfrak{U}$. Since any two charts in $\mathfrak{M} \setminus \mathfrak{U}$ are compatible with $\mathfrak{U}$, they are compatible with each other by Proposition 5.12. Thus $\mathfrak{M}$ is an atlas on M. Clearly $\mathfrak{U} \subset \mathfrak{M}$, and $\mathfrak{M}$ is maximal by construction. Now suppose $\mathfrak{M}'$ is another maximal atlas containing $\mathfrak{U}$. Then all the charts in $\mathfrak{M}'$ are compatible with $\mathfrak{U}$, and so $\mathfrak{M}' \subset \mathfrak{M}$ by construction of $\mathfrak{M}$. Since $\mathfrak{M}'$ is maximal by assumption, we have that $\mathfrak{M} = \mathfrak{M}'$. □

Remark 5.15. As a consequence of Theorem 5.14, to verify that M is a smooth manifold, it suffices to show that (i) M is a second countable Hausdorff space, and (ii) there exists an atlas $\mathfrak{U}$ (not necessarily maximal) on M. For then $\mathfrak{U}$ can be extended to a unique maximal atlas $\mathfrak{M}$ on M by the preceding theorem, making M into a smooth manifold. If M is a subset of any second countable Hausdorff space (e.g. $\mathbb{R}^n$) and is given the subspace topology, then it suffices only to construct an atlas on M since (i) and (ii) are automatic.

From now on, the word "manifold" will mean smooth manifold, unless explicitly stated to the contrary.

Example 5.16. The unit circle S^1 (as a subspace of $\mathbb{R}^2$) is a 1-manifold. As a set, $S^1 = \{(x, y) : x^2 + y^2 = 1\}$. According to Remark 5.15, we need only construct an atlas on it. Let U_1, U_2, U_3, and U_4, be the top, bottom, left, and right open semicircles respectively. For example, U_1 is the intersection of the open upper half plane $\{(x, y) : y > 0\}$ with S^1.

For $i = 1, 2, 3, 4$, define the coordinate maps $\phi_i : U_i \to (-1, 1) \subset \mathbb{R}$ as follows. For $i = 1, 2$, put $\phi_i(x, y) = x$. For $i = 3, 4$, put $\phi_i(x, y) = y$. Thus the top and bottom semicircles project onto the x-axis and the left and right semicircles project onto the y-axis. Each ϕ_i is a homeomorphism of U_i onto $(-1, 1)$. For example,

$$\phi_1 : U_1 \to (-1, 1) : (\xi, \sqrt{1 - \xi^2}) \mapsto \xi$$

is a continuous bijection of U_1 onto $(-1, 1)$, and it's inverse

$$\phi_1^{-1} : (-1, 1) \to U_1 : \xi \mapsto (\xi, \sqrt{1 - \xi^2})$$

is continuous as well. Hence it is a homeomorphism. It remains to check compatability of the coordinate charts. Consider the charts (U_2, ϕ_2) and (U_3, ϕ_3) for example. The other cases are handled similarly. We have

$$(\phi_2 \circ \phi_3^{-1})(\xi) = \phi_2(-\sqrt{1 - \xi^2}, \xi) = -\sqrt{1 - \xi^2},$$

which is C^∞ on $\phi_3(U_2 \cap U_3)$ and

$$(\phi_3 \circ \phi_2^{-1})(\xi) = \phi_3(\xi, -\sqrt{1 - \xi^2}) = -\sqrt{1 - \xi^2},$$

which is C^∞ on $\phi_2(U_2 \cap U_3)$.

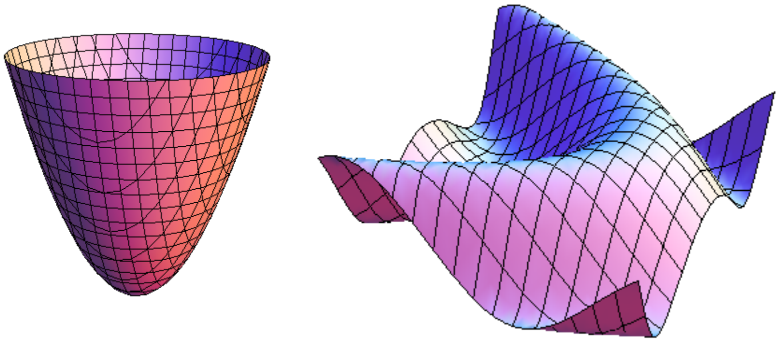
(*Note:* We could also cover S^1 with only *two* compatible charts by using the stereographic projection. This is a good exercise for the reader.)

Example 5.17. Let $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ be a C^∞ function with U open in $\mathbb{R}^n$. Then the *graph of* f, denoted $\Gamma(f)$ and defined by

$$\Gamma(f) = \{(x, f(x)) : x \in U\}$$

is an n-manifold. To see this, first identify $\mathbb{R}^n \times \mathbb{R}^m$ with $\mathbb{R}^{n+m}$. By Remark 5.15, it suffices to construct an atlas for $\Gamma(f)$. Indeed, define $\phi : \Gamma(f) \to U : (x, f(x)) \mapsto x$. Clearly ϕ is a continuous bijection, and $\phi^{-1} : U \to \Gamma(f) : x \mapsto (x, f(x))$ is continuous as well. Thus ϕ is a homeomorphism of $\Gamma(f)$ onto the open subset U of $\mathbb{R}^n$. Hence $\{(\Gamma(f), \phi)\}$ is an atlas (consisting of a single coordinate chart) convering $\Gamma(f)$, and so $\Gamma(f)$ is an n-manifold.

Consequently, many familiar surfaces from calculus are manifolds. For example, both the paraboloid $z = x^2 + y^2$ and the surface $z = \sin(x + y^2)$ are 2-manifolds. They are shown in in Figures 4(a) and 4(b), respectively.



(a) $z = x^2 + y^2$ (b) $z = \sin(x + y^2)$

Figure 5.4: Some simple 2-manifolds in $\mathbb{R}^3$.

The following Mathematica code produced the paraboloid $z = x^2 + y^2$.

```
ContourPlot3D[x^2+y^2==z, {x, -1, 1}, {y, -1, 1}, {z,0,1},
  Axes -> False, Boxed -> False]
```

This code produced the wavy surface $z = \sin(x + y^2)$.

```
Plot3D[Sin[x + y^2], {x, -3, 3}, {y, -2, 2}, Axes -> False,
  Boxed -> False]
```

Theorem 5.18. *Let N and M be manifolds. Then $N \times M$ is also a manifold when equipped with the product topology.*

Proof. Since $N \times M$ is given the product topology, it follows that $N \times M$ is Hausdorff and second countable, since these properties are preserved under finite products. Let $\{(U_\alpha, \phi_\alpha)\}$ and $\{(V_\beta, \psi_\beta)\}$ be atlases for N and M respectively. We claim that the collection $\mathcal{U}$ of all $(U_\alpha \times V_\beta, \phi_\alpha \times \psi_\beta)$ is an atlas for $N \times M$, where $\phi_\alpha \times \psi_\beta$ is a tensor-like product defined by

$$(\phi_\alpha \times \psi_\beta)(p, q) = \phi_\alpha(p)\psi_\beta(q).$$

First we show that if (U, ϕ) is a chart on N and (V, ψ) is a chart on M, then $(U \times V, \phi \times \psi)$ is a chart on $N \times M$. That $U \times V$ is open in $N \times M$ follows from the definition of the product topology on $N \times M$. Next, ϕ is a homeomorphism from U onto the open subset $\phi(U)$ of $\mathbb{R}^n$, and ψ is a homeomorphism from V onto the open subset of $\phi(V)$ of $\mathbb{R}^m$. It follows that $\phi \times \psi$ is a homeomorphism from the open subset $U \times V$ of $N \times M$ onto the open subset $\phi(U) \times \psi(V)$ of $\mathbb{R}^n \times \mathbb{R}^n \simeq \mathbb{R}^{n+m}$. In other words, the pair $(U \times V, \phi \times \psi)$ is a chart on $N \times M$.

Secondly, if (p, q) is any point in $N \times M$ then there is a chart (U_α, ϕ_α) in the atlas of N containing p and a chart (V_β, ψ_β) in the atlas of M containing q. Then $(U_\alpha \times V_\beta, \phi_\alpha \times \psi_\beta)$ is a chart containing (p, q). Hence $\mathcal{U}$ is a collection of charts on $N \times M$ that covers $N \times M$.

It remains to show that any two charts in $\mathcal{U}$ are compatible. Consider the charts $(U_\alpha \times V_\beta, \phi_\alpha \times \psi_\beta)$ and $(U_\gamma \times V_\delta, \phi_\delta \times \psi_\gamma)$ in $\mathcal{U}$. Then

$$\begin{aligned}(\phi_\gamma \times \psi_\delta) \circ (\phi_\alpha \times \psi_\beta)^{-1} &= (\phi_\gamma \times \psi_\delta) \circ (\phi_\alpha^{-1} \times \psi_\beta^{-1}) \\ &= (\phi_\gamma \circ \phi_\alpha^{-1}) \times (\psi_\delta \circ \psi_\beta^{-1}).\end{aligned}$$

Now, (U_γ, ϕ_γ) and (U_α, ϕ_α) are charts in the atlas of N, so they are compatible with each other. Thus $\phi_\gamma \circ \phi_\alpha^{-1}$ is C^∞ from $\phi_\alpha(U_\alpha \cap U_\gamma)$ onto $\phi_\gamma(U_\alpha \cap U_\gamma)$. Similarly, $\psi_\delta \circ \psi_\beta^{-1}$ is C^∞ from $\psi_\beta(V_\beta \cap V_\delta)$ onto $\psi_\delta(V_\beta \cap V_\delta)$. It follows that $(\phi_\gamma \times \psi_\delta) \circ (\phi_\alpha \times \psi_\beta)^{-1}$ is C^∞ on the open set $\phi_\alpha((U_\alpha \times V_\beta) \cap (U_\gamma \times V_\delta)) \times \psi_\beta((U_\alpha \times V_\beta) \cap (U_\gamma \times V_\delta))$. The proof that $(\phi_\alpha \times \psi_\beta) \circ (\phi_\gamma \times \psi_\delta)^{-1}$ is C^∞ is similar. Thus the two charts $(U_\alpha \times V_\beta, \phi_\alpha \times \psi_\beta)$ and $(U_\gamma \times V_\delta, \phi_\delta \times \psi_\delta)$ are compatible. Hence the collection $\mathcal{U}$ is indeed an atlas on $N \times M$. $\square$

Remark 5.19. Theorem 5.18 is a powerful theorem, for it enables us to easily construct more complicated manifolds from simpler ones. For example, the theorem implies that the torus $S^1 \times S^1$ and the cylinder $S^1 \times \mathbb{R}$ are manifolds. By induction we can continue this finitely many times. For example, the n-dimensional torus $S^1 \times S^1 \times \cdots \times S^1$ (with n factors) is an n-manifold.

5.3 Smooth Maps on a Manifold

Definition 5.20. Let M be an n-manifold. A function $f : M \to \mathbb{R}$ is said to be C^∞ (or *smooth*) *at a point* $p \in M$ if there exists a chart (U, ϕ) containing p in the atlas of M such that $f \circ \phi^{-1} : \phi(U) \to \mathbb{R}^n$ is C^∞ at $\phi(p)$.

Lemma 5.21. *A function $f : M \to \mathbb{R}$ is smooth at $p \in M$ iff for every chart (U, ϕ) containing p in the atlas of M we have $f \circ \phi^{-1} : \phi(U) \to \mathbb{R}^n$ is C^∞ at $\phi(p)$.*

Proof. Suppose f is smooth at p. Then there exists a chart (U, ϕ) containing p such that $f \circ \phi^{-1}$ is C^∞ at $\phi(p)$. Now let (V, ψ) be any other chart containing p in the atlas of M. Since these two charts are C^∞ compatible (because they're in the atlas of M), we have that $\phi \circ \psi^{-1} : \psi(U \cap V) \to \phi(U \cap V)$ is C^∞. Hence

$$f \circ \psi^{-1} = (f \circ \phi^{-1}) \circ (\phi \circ \psi^{-1})$$

is smooth at $\psi(p)$, since the composition of smooth functions is smooth. The converse is obvious. □

8 The Tangent Space

In chapter 2 we visualized a tangent vector at $p \in \mathbb{R}^n$ to be a geometrical arrow whose base point was p. However, this geometrical approach has at least two deficiencies, namely

(i) it does not lend itself to computation, and

(ii) it does not generalize well to manifolds, especially those that are not embeddable in a Euclidean space.

To remedy the first deficiency, we identified an arrow at p with coordinates in $\mathbb{R}^n$, thus turning the set of geometrical arrows attached to p into a vector space isomorphic with $\mathbb{R}^n$. We called this vector space the tangent space to $\mathbb{R}^n$ at p, and we denoted it by $T_p(\mathbb{R}^n)$. This made computation possible, but the second problem remained.

To address the second deficiency, we began by defining germs of functions and derivations at a point. We then established a somewhat striking result: the space $\mathcal{D}_p(\mathbb{R}^n)$ of all derivations at p is isomorphic to $T_p(\mathbb{R}^n)$. In symbols, $\mathcal{D}_p(\mathbb{R}^n) \simeq T_p(\mathbb{R}^n)$. Thus we can *identify* tangent vectors at p with derivations at p. This idea becomes the basis of our *definition* for tangent vectors at a point p in an arbitrary manifold M.

Indeed, if M is a manifold and $p \in M$, we will simply *define* a tangent vector at p to be a derivation at p. (Of course, we will show that we recover the usual definition in the special case of $\mathbb{R}^n$.) For this definition to make sense, we must first define a derivation at p, which in turn depends on the concept of germs of C^∞ functions at p. Thus we begin by defining germs.

But before we begin, one subtle technical point is worth bearing in mind. The definitions below will appear deceptively identical to their earlier counterparts in chapter 2. But there is a subtle difference. The difference lies in what it means for a function to be C^∞. In chapter 2 the manifold was $\mathbb{R}^n$ (and we know what it means for a function on $\mathbb{R}^n$ to be smooth). But for a function f on an arbitrary manifold M another definition is necessary. Indeed recall from chapter 6 that a function $f : M \to \mathbb{R}$ is C^∞ at $p \in M$ if there is a coordinate chart (U, ϕ) in the atlas of M containing p such that $f \circ \phi^{-1} : \phi(U) \to \mathbb{R}$ is C^∞ at $\phi(p)$. If U is a neighborhood of $p \in M$, we say $f : U \subset M \to \mathbb{R}$ is C^∞ if it is C^∞ at each $p \in U$.

Thus the following definitions are similar in language and spirit to those from chapter 2, but they are technically different; in particular they generalize those earlier definitions. Bearing all this in mind, we now define germs of functions at a point in an arbitrary manifold.

Definition 8.1. Let M be a manifold and $p \in M$. Let

$$\mathcal{S}_p = \{(f, U) : U \text{ is a neighborhood of } p \text{ and } f : U \to \mathbb{R} \text{ is } C^\infty\}.$$

Define an equivalence relation $\sim$ on $\mathcal{S}_p$ by saying that $(f, U) \sim (g, V)$ if and only if there exists a neighborhood $W \subset U \cap V$ containing p such that $f = g$ on W. The equivalence class of (f, U) is called the *germ* of f at p, and the set of all germs is denoted $C_p^\infty(M)$.

Remark 8.2. If we identify the pair (f, U) with just the function f, then loosely speaking we may regard the germ of f at p as the set of all functions that agree with f on some neighborhood of p. This means that if P is some property that depends only on f and its derivatives on a neighborhood of the point p, then either all functions in the germ of f satisfy property P or none of them do.

Definition 8.3. Let M be a manifold and $p \in M$. A *derivation* at p is a linear map $D : C_p^\infty(M) \to \mathbb{R}$ that satisfies the Leibniz rule, i.e.

$$D(fg) = f(p)(Dg) + g(p)(Df).$$

We also refer to D as a *point derivation* of $C_p^\infty(M)$.

We are now in a position to formally define a tangent vector and the tangent space at a point in an arbitrary manifold.

Definition 8.4. A *tangent vector* at a point p in a manifold M is a derivation at p in M. The set of all tangent vectors at the point p is called the *tangent space* of M at p and is denoted T_pM.

Remark 8.5.

1. As in chapter 2, let $\mathcal{D}_p(M)$ denote the set of all derivations at p in M. Then in fact $T_pM = \mathcal{D}_p(M)$ by *definition*. Contrast this to the situation in chapter 2 where $\mathcal{D}_p(\mathbb{R}^n)$ and $T_p(\mathbb{R}^n)$ were *different*, but isomorphic, vector spaces.

2. Define addition and scalar multiplication in T_pM in the usual way. Namely, if $a, b \in \mathbb{R}$, $X_p, Y_p \in T_pM$, and $f \in C_p^\infty(M)$, we define

   $$(aX_p + bY_p)f = a(X_pf) + b(Y_pf),$$

   and

   $$(aX_p)f = a(X_pf).$$

   Then it is easily seen that T_pM is a vector space.

3. By definition a tangent vector $X_p \in T_pM$ is really a special kind of *function*, namely a derivation at p in M. Referring to X_p as a "vector" is justified because we have just seen that T_pM is a vector space.

4. Consider a tangent vector to p in M, say $X_p \in T_pM$. Now X_p is in fact a derivation at p, and a derivation at p is a linear map that assigns a real number to each germ in $C_p^\infty(M)$. It follows that in order to fully describe the tangent vector X_p, we must specify its rule of assignment that takes germs of C^∞ functions into real numbers.

Now assume we have a C^∞ map $F : N \to M$ between manifolds. Associated to the point $p \in N$ is the tangent space T_pN, and associated to the point $F(p) \in M$ is the tangent space $T_{F(p)}M$. It seems reasonable that F should somehow induce a map F_* from T_pN to $T_{F(p)}M$, and that the definition of F_* should involve F in a natural way.

By part 4 of Remark 8.5, given $X_p \in T_pN$ the definition of F_* should describe how $F_*(X_p)$ acts on a function $f \in C_{F(p)}^\infty(M)$ to produce a real number. This is achieved by the next definition.

Definition 8.6. Let $F : N \to M$ be a C^∞ map between the manifolds N and M, and let $p \in N$. We define the differential of F at p, denoted $F_* : T_pN \to T_{F(p)}M$, as follows. For any $X_p \in T_pN$ and any $f \in C^\infty_{F(p)}(M)$, put

$$(F_*(X_p))f = X_p(f \circ F).$$

In other words, we pull f back by F and then operate on it by X_p. If we wish to emphasize the point p, we will write $F_{*,p}$ instead.

Remark 8.7.

1. For the definition to make sense, we must show that $f \circ F$ belongs $C^\infty_p(N)$ in order to operate on it with X_p. Strictly speaking f is a germ, but we may identify it with one of its representative C^∞ functions. With such an identification, the smoothness of f and the map F implies that $f \circ F$ is C^∞ on some neighborhood of $p \in N$. Hence $f \circ F$ belongs to some germ in $C^\infty_p(N)$, and if we identify $f \circ F$ with its germ then the expression $X_p(f \circ F)$ makes sense. Going forward we will assume such an identification when necessary without explicitly mentioning it.

2. We could write $(F_*(X_p))f = X_p(F^*f)$, where $F^*f = f \circ F$ denotes the pullback of f by F. But the notation is already confusing enough, and to use the $*$ symbol in two different ways in this equation would lead to apoplexy.

3. To help summarize and clarify the quantities involved:

   (a) X_p is a tangent vector in T_pN (i.e. X_p is a derivation at p).

   (b) F_* is a map from T_pN to $T_{F(p)}M$.

   (c) $F_*(X_p)$ is a tangent vector in $T_{F(p)}M$ (i.e. $F_*(X_p)$ is a derivation at $F(p)$). In particular, $F_*(X_p)$ is the derivation that sends $f \in C^\infty_{F(p)}(M)$ to the real number $X_p(f \circ F)$. We may write this as

   $$F_*(X_p) = (f \mapsto X_p(f \circ F)),$$

   or equivalently

   $$F_*(X_p) : C^\infty_{F(p)}(M) \to \mathbb{R} : f \mapsto X_p(f \circ F).$$

   (d) For any $f \in C^\infty_{F(p)}(M)$, the quantity $(F_*(X_p))f$ is a real number, and the rule of assignment is $(F_*(X_p))f = X_p(f \circ F)$.

(e) **Grand summary**: We can piece all the above information together into the following grand statement. Given manifolds N and M, a C^∞ map $F : N \to M$, and a point $p \in N$, the differential F_* at p is described as follows.

$$F_* : T_pN \to T_{F(p)}M : X_p \mapsto (f \mapsto X_p(f \circ F)), \quad \forall f \in C^\infty_{F(p)}(M).$$

If you can stare at that displayed equation and understand it, then move on.

Exercise 8.8. Verify that $F_* : T_pN \to T_{F(p)}M$ is a linear map and that $F_*(X_p)$ is a actually a derivation at $F(p)$.

Proof. To see that $F_* : T_pN \to T_{F(p)} \in M$ is a linear map, let $a, b \in \mathbb{R}$ and $X_p, Y_p \in T_pN$. Then for any $f \in C^\infty_{F(p)}(M)$, we have

$$\begin{aligned}(F_*(aX_p + bY_p))f &= (aX_p + bY_p)(f \circ F)\\ &= a(X_p(f \circ F)) + b(Y_p(f \circ F))\\ &= a((F_*(X_p))f) + b((F_*(Y_p))f).\end{aligned}$$

Hence $F_* : T_pN \to T_{F(p)}M$ is a linear map.

To see that $F_*(X_p)$ is a derivation at $C^\infty_{F(p)}(M)$, we must show that (i) it is a linear map from $C^\infty_{F(p)}(M)$ into $\mathbb{R}$, and (ii) it satisfies the Leibniz rule. Thus let $a, b \in \mathbb{R}$ and $f, g \in C^\infty_{F(p)}(M)$. Then

$$\begin{aligned}F_*(X_p)(af + bg) &= X_p((af + bg) \circ F)\\ &= X_p(a(f \circ F) + b(g \circ F))\\ &= aX_p(f \circ F) + bX_p(g \circ F)\\ &= aF_*(X_p)f + bF_*(X_p)g,\end{aligned}$$

where in going to the third equality I used the fact that X_p is a derivation, hence linear. This shows linearity of $F_*(X_p)$. For the Leibniz property,

$$\begin{aligned}F_*(X_p)(fg) &= X_p((fg) \circ F)\\ &= X_p((f \circ F)(g \circ F))\\ &= ((g \circ F)(p))X_p(f \circ F) + ((f \circ F)(p))X_p(g \circ F)\\ &= g(F(p))F_*(X_p)f + f(F(p))F_*(X_p)g,\end{aligned}$$

where in going to the third equality I used the Leibniz property of X_p. Thus $F_*(X_p)$ is linear and satisfies the Leibniz rule, hence it is a derivation as required. □

References

[1] William M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, second edition, 1986.

[2] John M. Lee. *Introduction to Smooth Manifolds*. Springer, 2003.

[3] James R. Munkres. *Topology*. Prentice Hall, second edition, 2000.

[4] Loring W. Tu. *An Introduction to Manifolds*. Springer, 2008.