BASIC DIFFERENTIABLE MANIFOLDS

TSOGTGEREL GANTUMUR

ABSTRACT. The most basic notions of differential geometry are reviewed.

CONTENTS

1. TOPOLOGICAL MANIFOLDS

Recall that a *topological space* is a set X, together with a prescription of which subsets of X are considered to be open. The collection of all open subsets of X is called the *topology* of X. More precisely, a set $\mathscr{T}$ of subsets of X is called a *topology* on X, if it satisfies the following axioms:

- $\varnothing \in \mathscr{T}$ and $X \in \mathscr{T}$ (The empty set and the space itself are open).
- If $A, B \in \mathscr{T}$ then $A \cap B \in \mathscr{T}$ (Finite intersection of open sets is open).
- If $\{A_\alpha\} \subset \mathscr{T}$ then $\cup_\alpha A_\alpha \in \mathscr{T}$ (Arbitrary union of open sets is open).

An example is given by $\mathbb{R}^n$ with its usual topology, that $A \subset \mathbb{R}^n$ is open iff for any $x \in A$ there is an open ball $B \ni x$ such that $B \subset A$. More generally, any metric space $(\mathscr{M}, \rho)$ becomes a topological space, when the open ball $B_r(x)$ centred at $x \in \mathscr{M}$ of radius $r > 0$ is defined as $B_r(x) = \{y \in \mathscr{M} : \rho(x, y) < r\}$.

A map $f : X \to Y$ between topological spaces is called *continuous* if for any open subset $B \subset Y$, the preimage $f^{-1}(B) = \{x \in X : f(x) \in B\}$ is open. For metric spaces, this notion reduces to the usual notion of continuity (Exercise). In addition to $f : X \to Y$ being continuous, if the inverse $f^{-1} : Y \to X$ exists and is continuous, then we say that f is a *homeomorphism* between X and Y, and that X and Y are *homeomorphic* to each other.

Definition 1.1. A *locally Euclidean space* (or an *LE-space*) is a topological space, which is locally homeomorphic to $\mathbb{R}^n$. More precisely, a topological space M is called an *n-dimensional locally Euclidean space*, if for any $p \in M$, there is an open set $U \subset M$ containing p, that is homeomorphic to an open subset $V \subset \mathbb{R}^n$. Each such pair (U, ϕ), where $\phi : U \to V$ is a homeomorphism, is called a *chart* of M, and a collection $\{(U_\alpha, \phi_\alpha)\}$ of charts satisfying the (covering) property $\cup_\alpha U_\alpha = M$ is called an *atlas* (or an *LE-structure*) on M. Thus a locally Euclidean space is a topological space equipped with an LE-structure.

Date: 2017/03/01 21:19 (EST). Incomplete draft.

Remark 1.2. In order to show that a given topological space X is an LE-space, it suffices to construct an atlas on X. Note that an atlas is *not* an additional structure on X, but a tool to reveal a certain property of X. This is akin to producing a nontrivial divisor to show that an integer m is a composite number.

Remark 1.3. Given an atlas $\{(U_\alpha, \phi_\alpha)\}$ on M, since each $\phi_\alpha : U_\alpha \to V_\alpha \subset \mathbb{R}^n$ is a homeomorphism, U_α is topologically identical to V_α, an open subset of $\mathbb{R}^n$. This gives a way to understand and work with M locally by using $\phi_\alpha : U_\alpha \to \mathbb{R}^n$ as a *coordinate system* on U_α. Hence in a certain sense, an LE-space M itself is nothing more than a prescription of how the sets $V_\alpha \subset \mathbb{R}^n$ are glued together.

Simple examples of locally Euclidean spaces are $\mathbb{R}^n$ and any open subset of $\mathbb{R}^n$. Recall that given a topological space X and a subset $Y \subset X$, the *subspace topology* on Y is defined by saying that $B \subset Y$ is open iff there is an open $A \subset X$ such that $B = A \cap Y$. With the subspace topology, any linear (or affine) subspace of $\mathbb{R}^n$ is a locally Euclidean space, since it would be homeomorphic to $\mathbb{R}^k$ for some k.

Example 1.4. Let $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, and let the maps $\pi_i : S^1 \to \mathbb{R}$ for $i = 1, 2$ be defined by $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$. Define $U_i \subset S^1$, $i = 1, 2, 3, 4$, as

$$U_1 = S^1 \cap \{y > 0\}, \quad U_2 = S^1 \cap \{x > 0\}, \quad U_3 = S^1 \cap \{y < 0\}, \quad U_4 = S^1 \cap \{x < 0\}. \tag{1}$$

By construction, these sets are open as subsets of S^1, and they cover S^1. In order to complete $\{U_i\}$ into an atlas, we want to show that each U_i is homeomorphic to an open interval of $\mathbb{R}$. Since any $p \in S^1$ is in some U_i, this would show that S^1 is a 1-dimensional topological manifold. To this end, let the bijections $\phi_i : U_i \to (-1, 1) \subset \mathbb{R}$ be given by

$$\phi_1 = \pi_1|_{U_1}, \quad \phi_2 = \pi_2|_{U_2}, \quad \phi_3 = \pi_1|_{U_3}, \quad \phi_4 = \pi_2|_{U_4}. \tag{2}$$

These maps are in fact homeomorphisms. We will demonstrate it only for ϕ_1. The preimage of an open set $B \subset (-1, 1)$ under ϕ_1 is $\phi_1^{-1}(B) = U_1 \cap (B \times \mathbb{R})$, which is obviously open in S^1. Hence ϕ_1 is continuous. To show that $\phi_1^{-1} : (-1, 1) \to U_1$ is continuous, let $A \subset U_1$ be such that $A = U_1 \cap D$, where $D \subset \mathbb{R}^2$ is an open disk. Then A can be described as the arc of U_1 bounded by two points $(x_1, y_1) \in U_1$ and $(x_2, y_2) \in U_1$, without including the endpoints. Now it is clear that either $\phi_1(A) = (x_1, x_2)$ or $\phi_1(A) = (x_2, x_1)$, and since any open set $O \subset U_1$ can be written as the union of such open arcs, we conclude that $\phi_1(O)$ is open in $(-1, 1)$, and thus $\phi_1^{-1} : (-1, 1) \to U_1$ is continuous. We conclude that $\{(U_i, \phi_i) : i = 1, 2, 3, 4\}$ is an atlas on S^1, and hence S^1 is a 1-dimensional LE-space.

Example 1.5. In the preceding example, the locally Euclidean space S^1 was realized as a subset of $\mathbb{R}^2$. Here we will construct it without reference to any ambient space. We define $\tilde{S}^1 = [0, 1)$ as a set, and introduce a topology on it by saying that $A \subset \tilde{S}^1$ is open iff it is the union of sets of the form either (a, b) or $[0, a) \cup (b, 1)$, with $0 < a < b < 1$. Intuitively speaking, we "identify" the endpoints of the interval $[0, 1]$. The same effect can be achieved by letting $\tilde{S}^1 = [0, 1]/\sim$, where $\sim$ is the equivalence relation defined on $[0, 1]$ by $x \sim y$ iff $x = y$ or $x, y \in \{0, 1\}$, and by equipping $\tilde{S}^1$ with the *quotient topology*. The latter means that $A \subset \tilde{S}^1$ is open iff the preimage of A under the map $x \mapsto [x]$, that sends x to the equivalence class containing it, is open in $[0, 1]$. Finally, to show that $\tilde{S}^1$ is homeomorphic to S^1, we define the bijection $f : \tilde{S}^1 \to S^1$ by $f(t) = (\cos 2\pi t, \sin 2\pi t)$. The topology on S^1 is generated by the open arcs, that are exactly the images under f of sets of the form either (a, b) or $[0, a) \cup (b, 1)$, and therefore f is a homeomorphism.

Example 1.6. If we wanted to show directly that $\tilde{S}^1$ as in the preceding example is an LE-space, an LE-structure on $\tilde{S}^1$ can be introduced as follows. Let $U_1 = (0, 1)$ and $U_2 =$

$[0,\frac{1}{2})\cup(\frac{1}{2},1)$, which are both open in $\tilde{S}^1$. Define the maps $\phi_1 : U_1 \to V_1 = (0,1)$ by $\phi_1(x) = x$, and $\phi_2 : U_2 \to V_2 = (\frac{1}{2},\frac{3}{2})$ by

$$\phi_2(x) = \begin{cases} x+1 & \text{for } 0 \leq x < \frac{1}{2}, \\ x & \text{for } \frac{1}{2} < x < 1. \end{cases} \tag{3}$$

These maps are clearly homeomorphisms, and so $\{(U_1,\phi_1),(U_2,\phi_2)\}$ is an atlas on $\tilde{S}^1$. Both coordinate maps ϕ_1 and ϕ_2 are defined on $U_1 \cap U_2 = \tilde{S}^1 \setminus \{0,\frac{1}{2}\}$, meaning that the points in $U_1 \cap U_2$ will have coordinate representations in both V_1 and V_2. The two different coordinates of such points are related by the so-called *transition maps*

$$\phi_{12} = \phi_1 \circ \phi_2^{-1} : \phi_2(U_1 \cap U_2) \to \phi_1(U_1 \cap U_2), \tag{4}$$

and $\phi_{21} = \phi_2 \circ \phi_1^{-1} = \phi_{12}^{-1}$. We can explicitly compute

$$\begin{aligned} \phi_1(U_1 \cap U_2) = (0,\tfrac{1}{2}) \cup (\tfrac{1}{2},1), \qquad \phi_2(U_1 \cap U_2) = (\tfrac{1}{2},1) \cup (1,\tfrac{3}{2}), \\ \phi_{12}(t) = \begin{cases} t & \text{for } \frac{1}{2} < t < 1, \\ t-1 & \text{for } 0 < t < \frac{1}{2}. \end{cases} \end{aligned} \tag{5}$$

Now, the point we wanted to make is that $\tilde{S}^1$ is completely determined by V_1, V_2, and the transition maps. To see this, define the equivalence relation $\sim$ on the disjoint union $V_1 \sqcup V_2$ by $V_1 \ni t_1 \sim t_2 \in V_2$ iff $t_1 = \phi_{12}(t_2)$. Then $\tilde{S}^1$ is homeomorphic to $V_1 \sqcup V_2/\sim$. Basically, V_1 and V_2 are the coordinate systems or the pieces of $\mathbb{R}$ to be used in the construction of $\tilde{S}^1$, and the transition maps specify how the pieces are glued together. What this construction allows for, as opposed to, e.g., identifying the endpoints of the interval $[0,1]$ as in the preceding example, is that the "overlapping" regions are "wider," and therefore, as we will see, information such as how to differentiate functions can be exchanged between V_1 and V_2.

Now we make a new definition to rule out some pathological examples of LE-spaces.

Definition 1.7. A *topological manifold* is a locally Euclidean space, that is Hausdorff and second countable.

Remark 1.8. An alternative, more relaxed definition can be obtained by replacing the second countability condition with paracompactness. However, the difference is rather artificial, in that a paracompact Hausdorff LE-space is *not* second countable if and only if it has uncountably many connected components. In particular, each connected component of a paracompact Hausdorff LE-space is a topological manifold in the sense of Definition 1.7.

Recall that a topological space X is *Hausdorff* if for any $x,y \in X$ with $x \neq y$, there exist disjoint open sets $U \ni x$ and $V \ni y$. In other words, in a Hausdorff space, points can be separated by open sets. On the other hand, a topological space X is *second countable* if there exists a countable collection $\mathscr{B}$ of open subsets of X, such that every open subset $O \subset X$ can be written as $O = \cup_{A\in\mathscr{A}} A$ for some $\mathscr{A} \subset \mathscr{B}$. A basic example of a second countable space is $\mathbb{R}^n$, since we can take $\mathscr{B}$ in this case to be the collection of open balls with rational radii and centres at points with rational coordinates. Thus the second countability condition prevents topological manifolds from being "as large as uncountably many $\mathbb{R}^n$'s." An example of a non-second-countable LE-space would be the long line.

Example 1.9. We have seen that open subsets of $\mathbb{R}^n$, linear subspaces of $\mathbb{R}^n$, and the circle S^1 are LE-spaces. Since each of these examples can be considered as a topological subspace of $\mathbb{R}^n$, and as such inherits the Hausdorff and second countability properties of $\mathbb{R}^n$, they are actually topological manifolds.

Example 1.10. The Hausdorffness condition is very reasonable and covers most LE-spaces that come up in practice, but one does not need to look far in order to find a non-Hausdorff LE-space. In the xy-plane, consider the dynamical system

$$\begin{cases} x' = 0, \\ y' = |x| + |y|. \end{cases} \tag{6}$$

All solution curves have $x(t) = c = \text{const}$. For $c = 0$, we have 3 curves: $y(t) = 0$, $y(t) = e^t$, and $y(t) = -e^{-t}$. For each $c \neq 0$, we have the single solution curve $y(t) = |c| \operatorname{sgn}(t)(e^{|t|} - 1)$. Let us define an equivalence relation on $\mathbb{R}^2$ by $p \sim q$ iff $p \in \mathbb{R}^2$ and $q \in \mathbb{R}^2$ are on the same solution curve, and let $X = \mathbb{R}^2/\!\sim$, that is, X is the space of solution curves. We can introduce an LE-structure on X with the help of the atlas $\{(U_i, \phi_i) : i = 0, \pm 1\}$, where $U_i = \{[(x, i)] : x \in \mathbb{R}\}$ and $\phi_i([(x, i)]) = x$, where $[(x, i)]$ is the solution curve containing the point (x, i).

Exercise 1.1. In a metric space X with metric ρ, we define the *ball* centered at $x \in X$, of radius r, to be the set $B_r(x) \equiv B(x, r) = \{y \in X : \rho(x, y) < r\}$. A subset $S \subset X$ is called *open* if for any $x \in S$, there is $\varepsilon > 0$ such that $B_\varepsilon(x) \subset S$. Show that the collection of all open sets of X forms a topology on X.

Exercise 1.2. We say that a sequence $\{x_n\} \subset X$ *converges* to $x \in X$, and write $x_n \to x$, if for any open $U \subset X$ with $U \ni x$, there is an index m such that $x_n \in U$ for all $n > m$. A subset $B \subset X$ is said to be *closed* if $\{x_n\} \subset B$ and $x_n \to x \in X$ imply that $x \in B$. Show that $B \subset X$ is closed if and only if its complement $X \setminus B$ is open.

Exercise 1.3. Let $f : X \to Y$ be a map between two metric spaces. Show that the following are equivalent.

(a) f is continuous.
(b) Whenever $U \subset Y$ is open, its preimage $f^{-1}(U) = \{x \in X : f(x) \in U\}$ is open.
(c) The preimage of any closed $U \subset Y$ is closed.

Exercise 1.4. Show that every LE-space has the following properties.

(a) It is first countable and satisfies the separation axiom T_1.
(b) Its connected components are LE-spaces.
(c) It is connected if and only if it is path-connected.

Exercise 1.5. Show that every topological manifold is separable, σ-compact, Lindelöf, and paracompact.

Exercise 1.6. Let $\Omega \subset \mathbb{R}^n$ be open, and let $f : \Omega \to \mathbb{R}$ be a continuous function. Show that the graph $\Gamma = \{(x, f(x)) : x \in \Omega\}$, as a subset of $\mathbb{R}^{n+1}$, is a topological manifold.

Exercise 1.7. Show that $S^{n-1} = \{x \in \mathbb{R}^n : x_1^2 + \ldots + x_n^2 = 1\}$ is a compact topological manifold. Show that any atlas on S^{n-1} must include at least two charts.

2. Differentiable manifolds

Since a topological manifold M is also a topological space, the notions of continuous functions $f : M \to \mathbb{R}$ and continuous maps $f : M \to X$ (for a topological space X) are well defined. Given a continuous function $f : M \to \mathbb{R}$ and any chart (U, ϕ) of M, the composition $f \circ \phi^{-1}$ is a function on $\phi(U) \subset \mathbb{R}^n$, and thus we can check if $f \circ \phi$ is differentiable. However, if we want to use these compositions to introduce a notion of differentiable or smooth functions on M, then we are forced to work with only a *subset* of all possible charts (or atlases) on M, because one function can be differentiable in one chart but not differentiable in another, depending on the peculiarities of the coordinate systems. In order to have a consistent definition of differentiability, we need to work with an atlas whose transition maps are differentiable. These considerations lead to the following definition.

Definition 2.1. Given a topological manifold M and an integer $0 \leq k \leq \infty$, an atlas of M whose transition maps are $\mathscr{C}^k$ is called a $\mathscr{C}^k$ *structure* on M. That is, a $\mathscr{C}^k$ *structure* on M is an atlas $\{(U_\alpha, \phi_\alpha)\}$ of M, such that $\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(U_\alpha \cap U_\beta) \to \phi_\alpha(U_\alpha \cap U_\beta)$ is a $\mathscr{C}^k$ map for all α and β. Then a $\mathscr{C}^k$ *manifold* is a topological manifold equipped with a $\mathscr{C}^k$ structure. In case $k = \infty$, instead of the symbol $\mathscr{C}^\infty$ we typically use the word *smooth*. Smooth manifolds are often called simply *manifolds*.

If M is a $\mathscr{C}^k$ manifold, and if $f : M \to \mathbb{R}$ is such that $f \circ \phi^{-1} : \phi(U) \subset \mathbb{R}^n \to \mathbb{R}$ is $\mathscr{C}^k$ for some chart (U, ϕ) from the $\mathscr{C}^k$ structure of M, then $(f \circ \psi^{-1})|_{\psi(W \cap U)}$ is also $\mathscr{C}^k$ for any chart (W, ψ) from the $\mathscr{C}^k$ structure, because $(f \circ \psi^{-1})|_{\psi(W\cap U)} = (f \circ \phi^{-1})|_{\phi(W \cap U)} \circ (\phi \circ \psi^{-1})|_{\psi(W\cap U)}$. Thus the notion of $\mathscr{C}^k$ functions makes sense on a $\mathscr{C}^k$ manifold. More generally, we can define $\mathscr{C}^k$ maps between $\mathscr{C}^k$ manifolds, as follows.

Definition 2.2. A map $f : M \to N$ between two $\mathscr{C}^k$ manifolds is called a $\mathscr{C}^k$ *map* if the map $\psi|_{f(U)} \circ f|_U \circ \phi^{-1} : \phi(U) \to \psi(W)$ is a $\mathscr{C}^k$ map whenever (U, ϕ) and (W, ψ) are charts of M and N, respectively, with $W \cap f(U) \neq \varnothing$. If in addition, $f^{-1} : N \to M$ exists and is a $\mathscr{C}^k$ map, then f is called a $\mathscr{C}^k$*-diffeomorphism*, and M is said to be $\mathscr{C}^k$*-diffeomorphic* to N. For $k = \infty$, we omit the symbol $\mathscr{C}^\infty$, and simply talk about diffeomorphisms and diffeomorphic manifolds.

Some remarks and examples are in order.

Remark 2.3. Note that a $\mathscr{C}^0$ manifold is simply a topological manifold, and a $\mathscr{C}^k$ manifold is also a $\mathscr{C}^\ell$ manifold for each $\ell < k$.

Remark 2.4. By replacing the symbol $\mathscr{C}^k$ by "real analytic" or $\mathscr{C}^\omega$ in the preceding definitions, we get the definitions of *real analytic (or $\mathscr{C}^\omega$) manifolds and maps*.

Remark 2.5. Let M be a $\mathscr{C}^k$ manifold, and let $\{(U_\alpha, \phi_\alpha)\}$ be its $\mathscr{C}^k$ structure. Suppose that (U, ϕ) is an additional chart on M such that the transition maps $\phi \circ \phi_\alpha^{-1}$ and $\phi_\alpha \circ \phi^{-1}$ are $\mathscr{C}^k$ for all α. Then the new atlas $\{(U_\alpha, \phi_\alpha)\} \cup \{(U, \phi)\}$ is also a $\mathscr{C}^k$ structure on M, and replacing the old atlas by the new atlas does *not* change the class of $\mathscr{C}^k$ functions on M, and the classes of $\mathscr{C}^k$ maps to- and from M, given any fixed $\mathscr{C}^k$ manifold as the domain or the target. Therefore we should consider these two atlases as equivalent $\mathscr{C}^k$ structures. More formally, a chart (U, ϕ) is called $\mathscr{C}^k$*-compatible* with a given $\mathscr{C}^k$ structure $\mathscr{A}$ if the union $\mathscr{A} \cup \{(U, \phi)\}$ is also a $\mathscr{C}^k$ structure. Furthermore, two $\mathscr{C}^k$ structures are called $\mathscr{C}^k$*-equivalent* or $\mathscr{C}^k$*-compatible*, if their union is also a $\mathscr{C}^k$ structure. Then we *redefine* a $\mathscr{C}^k$ structure to be a $\mathscr{C}^k$-equivalence class of atlases. The same effect can be obtained by defining a $\mathscr{C}^k$ structure as the *maximal atlas* that is $\mathscr{C}^k$-compatible to a given one, i.e., as the atlas consisting of all charts that are $\mathscr{C}^k$-compatible with a given atlas. In practice, all this basically means that we still describe a $\mathscr{C}^k$ structure by a particular atlas, but implicitly assume that we are including all possible charts that are $\mathscr{C}^k$-compatible with that atlas.

Example 2.6. The natural equivalence relation between $\mathscr{C}^k$ manifolds is of course given by $\mathscr{C}^k$-diffeomorphisms. If no relevant additional structure is present, two $\mathscr{C}^k$ manifolds that are $\mathscr{C}^k$-diffeomorphic to each other must be considered identical. This notion of diffeomorphism equivalence is more relaxed than the $\mathscr{C}^k$-equivalence between atlases. As an example, take $M = \mathbb{R}$ with the smooth structure induced by the chart $\phi(x) = x$, and $N = \mathbb{R}$ with the smooth structure induced by the chart $\psi(x) = \sqrt[3]{x}$. Of course, M is simply $\mathbb{R}$ with its usual smooth structure. Thinking of M and N as identical topological manifolds, the charts (ϕ, M) and (ψ, M) are not even $\mathscr{C}^1$-compatible, because the transition map $(\psi \circ \phi^{-1})(t) = \sqrt[3]{t}$ is not differentiable at $t = 0$. So the maximal atlas containing (ϕ, M) is different from the one containing (ψ, M), meaning that M and N are different manifolds, even though they are identical as topological manifolds. However, the map $F : M \to N$ defined by $F(x) = x^3$ is a diffeomorphism, since $(\psi \circ F \circ \phi^{-1})(t) = t$ and $(\phi \circ F^{-1} \circ \psi^{-1})(t) = t$. Thus the smooth structure of N can be described by saying that $g : N \to \mathbb{R}$ is smooth if and only if $g \circ F$ is smooth on M. For instance, $g(x) = \sqrt[3]{x}$ is smooth on N. It might give the impression that N has a large supply of smooth functions as compared to M, but it is not true. The smooth functions on N are simply the usual smooth functions "twisted" by the map F, and the smooth structures of M and of N are identical up to a diffeomorphism.

Remark 2.7. The following questions are fundamental.

- Given a $\mathscr{C}^k$ manifold M, is there a $\mathscr{C}^\ell$ structure on M with $\ell > k$ that is $\mathscr{C}^k$-equivalent to the existing $\mathscr{C}^k$ structure?
- If the preceding question has an affirmative answer, is the $\mathscr{C}^\ell$ structure unique (up to $\mathscr{C}^\ell$-diffeomorphisms)?

For $1 \leq k < \ell \leq \infty$, both questions were answered in the affirmative by Whitney (1936). Moreover, he showed that different $\mathscr{C}^k$ structures give rise to different $\mathscr{C}^\ell$ structures. So the questions reduce to the case $k = 0$ and $\ell = \infty$. For dimensions $n = 1, 2, 3$, every topological manifold admits a unique smooth structure (Radó 1925, Moise 1952). On the other hand, by the works of many mathematicians including Milnor, Kervaire, Kirby, Siebenmann, Freedman, and Donaldson, accomplished during the period 1950's to 1980's, it is known that both answers are negative for $n \geq 4$ and $k = 0$. Namely, for each $n \geq 4$, there exists an n-dimensional topological manifold which does not admit any smooth structure, and there exists an n-dimensional topological manifold which admit more than one smooth structures that are not diffeomorphic to each other.

The following is a very useful criterion to recognize if the zero set of a function is a manifold. We invite the reader to prove it by using the implicit function theorem, and also to look up on the more general *constant rank theorem*.

Theorem 2.8 (Preimage theorem)**.** *Let $A \subset \mathbb{R}^n$ be an open set, and let $\phi : A \to \mathbb{R}^\ell$ be a $\mathscr{C}^k$ function, with $1 \leq k \leq \infty$. Suppose that for each $x \in A$ satisfying $\phi(x) = 0$, the map $D\phi(x) : \mathbb{R}^n \to \mathbb{R}^\ell$ is surjective. Then the set $M = \{x \in A : \phi(x) = 0\}$ is an $(n-\ell)$-dimensional $\mathscr{C}^k$ manifold.*

If $\ell = 1$ in the preceding theorem, then $D\phi(x)$ is a $1 \times n$ matrix, and surjectivity of $D\phi(x)$ simply means that $D\phi(x)$ has a nonzero entry. This special case is important enough to deserve a separate display.

Corollary 2.9 (Level surface theorem)**.** *Let $A \subset \mathbb{R}^n$ be an open set, and let $\phi : A \to \mathbb{R}$ be a $\mathscr{C}^k$ function, with $1 \leq k \leq \infty$. Suppose that $D\phi(x) \neq 0$ whenever $x \in A$ satisfies $\phi(x) = 0$. Then the set $M = \{x \in A : \phi(x) = 0\}$ is a $\mathscr{C}^k$ hypersurface in $\mathbb{R}^n$.*

Example 2.10 (Generalized sphere)**.** Let $a \in \mathbb{R}^n$ be a nonzero vector, and let

$$M = \{x \in \mathbb{R}^n : a_1 x_1^2 + a_2 x_2^2 + \ldots + a_n x_n^2 = 1\}. \tag{7}$$

We would like to show that M is a hypersurface. Thus we let

$$\phi(x) = a_1 x_1^2 + a_2 x_2^2 + \ldots + a_n x_n^2 - 1, \tag{8}$$

so that $M = \{\phi = 0\}$, and compute

$$D\phi(x) = (2a_1 x_1, 2a_2 x_2, \ldots, 2a_n x_n). \tag{9}$$

Since a is a nonzero vector, $D\phi(x) = 0$ if and only if $x = 0$. We know that $0 \notin M$, because $\phi(0) = -1$, and hence $D\phi(x) \neq 0$ for all $x \in M$. Then the level surface theorem implies that M is a smooth hypersurface in $\mathbb{R}^n$.

Example 2.11 (Orthogonal group)**.** Consider the set

$$M = \{X \in \mathbb{R}^{n\times n} : X^T X = I\}, \tag{10}$$

which is called the group of orthogonal matrices. This can be written as the zero set of $\phi : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, which is given by

$$\phi(X) = X^T X - I. \tag{11}$$

Although ϕ sends $n \times n$ matrices to $n \times n$ matrices, the output $\phi(X)$ has fewer than n^2 independent components, because $\phi(X)$ is always a symmetric matrix. Thus we think of ϕ as a mapping $\phi : \mathbb{R}^N \to \mathbb{R}^k$, with $N = n^2$ and $k = \frac{1}{2}n(n+1)$. In view of Theorem 2.8, our first task is to compute the directional derivative of ϕ along a matrix $B \in \mathbb{R}^{n\times n}$. Let us denote the components of ϕ, X, and B by ϕ_{ij}, x_{lm}, and b_{lm}, respectively. Then we have

$$\frac{\partial \phi_{ij}}{\partial x_{lm}}(X) = \frac{\partial}{\partial x_{lm}} \sum_{q=1}^{n} x_{qi}x_{qj} = \sum_{q=1}^{n} (\delta_{ql}\delta_{im}x_{qj} + x_{qi}\delta_{ql}\delta_{jm}) = \delta_{im}x_{lj} + x_{li}\delta_{jm}, \tag{12}$$

for the partial derivatives, and

$$D_B\phi_{ij}(X) = \sum_{l,m=1}^{n} (\delta_{im}x_{lj} + x_{li}\delta_{jm})b_{lm} = \sum_{l=1}^{n} (x_{lj}b_{li} + x_{li}b_{lj}) = (B^T X + X^T B)_{ij}, \tag{13}$$

for the directional derivative, yielding

$$D_B\phi(X) = X^T B + B^T X. \tag{14}$$

Our next task is to show that for each $X \in M$ and for any symmetric matrix $S \in \mathbb{R}^{n\times n}$, there exists $B \in \mathbb{R}^{n\times n}$ such that $D_B\phi(X) = S$. This would guarantee that $D\phi(X)$, as a linear map sending $\mathbb{R}^{n\times n}$ into the space of symmetric $n \times n$ matrices, is surjective. Let $S \in \mathbb{R}^{n\times n}$ be a symmetric matrix. We observe that $(X^T B)^T = B^T X$, and so the equation $X^T B + B^T X = S$ is of the form $C + C^T = S$. It is not difficult to construct a matrix C satisfying $C + C^T = S$. For example, one can check that the following works.

$$C_{ij} = \begin{cases} s_{ij} & \text{for } i < j, \\ \frac{1}{2}s_{ii} & \text{for } i = j, \\ 0 & \text{for } i > j. \end{cases} \tag{15}$$

Now that we have C, we need to solve $X^T B = C$. At this point, we recall that $X \in M$, that is, $X^T X = I$. This means that $(X^T)^{-1} = X$, and hence $B = XX^T B = XC$. We can also independently check that

$$X^T B + B^T X = X^T XC + (XC)^T X = C + C^T X^T X = C + C^T = S. \tag{16}$$

We conclude that the orthogonal group $M = \{X \in \mathbb{R}^{n\times n} : X^T X = I\}$ is a smooth manifold of dimension $N - k = \frac{1}{2}n(n-1)$. The standard notation for this manifold is $O(n) = M$ (not to be confused with the big-O notation).

Remark 2.12. In the preimage theorem and in the preceding examples, smooth manifolds were realized as subsets of $\mathbb{R}^n$. This situation is the easiest to visualize, and moreover, it turns out no loss of generality to study such manifolds, as the *Whitney embedding theorem* guarantees that any m-dimensional smooth manifold can be realized as a subset of $\mathbb{R}^{2m}$.

Now we introduce an extremely important tool in the study of manifolds. A collection $\{U_\alpha\}$ of subsets of a manifold M is called a *cover of* M if $M = \bigcup_\alpha U_\alpha$, and it is called an *open cover* if each U_α is open. We say that a cover $\{U_\alpha\}$ of M is *locally finite* if for each $x \in M$ there is an open set O containing x, such that O intersects with only finitely many U_α. Furthermore, a cover $\{V_i\}$ of M is called a *refinement of* $\{U_\alpha\}$ if for each V_i there is U_α

such that $V_i \subset U_\alpha$. We also recall the notation $B_r = \{x \in \mathbb{R}^n : |x| < r\}$ for the n-dimensional open ball of radius $r > 0$. The following theorem guarantees that any open cover of M has a refinement that is an atlas and has certain nice properties. This gives a lot of freedom in choosing coordinate charts on any given manifold.

Theorem 2.13. *Let M be a $\mathscr{C}^k$ manifold, and let $\{U_\alpha\}$ be an open cover of M. Then there exists a countable collection $\{V_i\}$ of subsets of M, satisfying the following properties.*

- *V_i is open, and $\overline{V}_i$ is compact for each i.*
- *$\{V_i\}$ is a locally finite refinement of $\{U_\alpha\}$.*
- *For each i, there exists a $\mathscr{C}^k$-diffeomorphism $\phi_i : V_i \to B_1$.*
- *$\{\phi_i^{-1}(B_{1/2})\}$ is an open cover of M.*

As an immediate corollary, we can prove the so-called "partition of unity property" for manifolds. Recall that the *support* of a function $f : M \to \mathbb{R}$, denoted by $\mathrm{supp} f$, is the closure of the set $M \setminus f^{-1}(\{0\}) = \{x \in M : f(x) \neq 0\}$.

Corollary 2.14 (Partition of unity)**.** *Let M be a $\mathscr{C}^k$ manifold, and let $\{U_\alpha\}$ be an open cover of M. Then there exists a countable collection $\{f_i\}$ of $\mathscr{C}^k$ functions on M, satisfying the following properties.*

- *$0 \leq f_i \leq 1$, and $\mathrm{supp} f_i$ is compact for each i.*
- *$\{\mathrm{supp} f_i\}$ is a locally finite refinement of $\{U_\alpha\}$.*
- *$\sum_i f_i(x) = 1$ for each $x \in M$.*

The collection $\{f_i\}$ as above is called a *$\mathscr{C}^k$ partition of unity subordinate to* $\{U_\alpha\}$. We will not prove these results, but advise the reader to find and read a proof elsewhere.

Remark 2.15. In fact, it can be argued that the partition of unity property given by the preceding theorem is as fundamental as the Hausdorff and the second countability (or paracompactness) properties, and therefore could even be included in the definition of manifolds, cf. Exercise 2.6 below. The importance of the partition of unity property can be illustrated by the following partial list of some of its applications and consequences.

- Definition of integration over manifolds.
- Existence of a Riemannian metric on any manifold.
- Existence of a volume form on any orientable manifold.
- The Whitney embedding theorem.

Exercise 2.1. Let $\Omega \subset \mathbb{R}^n$ be open, and let $f : \Omega \to \mathbb{R}$ be a $\mathscr{C}^k$ function. Show that the graph $\Gamma = \{(x, f(x)) : x \in \Omega\}$, as a subset of $\mathbb{R}^{n+1}$, is a $\mathscr{C}^k$ manifold.

Exercise 2.2. Show that $M = \{(x, y, z) \in \mathbb{R}^3 : \phi(x, y, z) = 1\}$ is a smooth manifold, where

$$\phi(x, y, z) = (\sqrt{x^2 + y^2} - 2)^2 + z^2. \tag{17}$$

Do you recognize this surface?

Exercise 2.3. Let $A \in \mathbb{R}^{n\times n}$ be an invertible, symmetric matrix. Then in each of the following cases, show that M is a smooth manifold, and determine the dimension of M.

(a) $M = \{x \in \mathbb{R}^n : x^T Ax = 1\} \subset \mathbb{R}^n$.
(b) $M = \{X \in \mathbb{R}^{n\times n} : X^T AX = A\} \subset \mathbb{R}^{n\times n}$.

Exercise 2.4. Let $O(n)$ be as in Example 2.11, and let $SO(n) = O(n) \cap \{\det A = 1\}$. Show that $O(n)$ and $SO(n)$ are compact. Identify the connected components of $O(n)$. Show that $SO(2)$ is diffeomorphic to S^1.

Exercise 2.5. Show that the Grassmanian

$$G(k,\mathbb{R}^n) = \{V \subset \mathbb{R}^n : V \text{ is a } k\text{-dimensional linear subspace of } \mathbb{R}^n\}, \tag{18}$$

is a compact smooth manifold.

Exercise 2.6. Let X be an LE-space that admits a $\mathscr{C}^0$ partition of unity subordinate to any open cover. Show that X is Hausdorff, paracompact and metrizable.

Exercise 2.7. Show that $\mathbb{R}^n$ and $\mathbb{R}^k$ are diffeomorphic to each other iff $n = k$.

Exercise 2.8. Prove that any smooth, connected 1-dimensional manifold is diffeomorphic to S^1 or to an open interval of $\mathbb{R}$.

3. Tangent spaces

If $\gamma : \mathbb{R} \to \mathbb{R}^m$ is a smooth function, understood as a parameterized curve, then the *velocity vector* at the parameter value $t \in \mathbb{R}$ is $\gamma'(t) \in \mathbb{R}^m$. Now, if $M \subset \mathbb{R}^m$ is a manifold, then its *tangent space* T_pM at $p \in M$ is the collection of all velocity vectors $\gamma'(0)$ as γ varies over all possible curves $\gamma : \mathbb{R} \to M$ with $\gamma(0) = p$. Namely, we define

$$T_pM = \{\gamma'(0) : \gamma \in \mathscr{C}^\infty(\mathbb{R},\mathbb{R}^m),\, \gamma(\mathbb{R}) \subset M,\, \gamma(0) = p\}. \tag{19}$$

Here, the fact that M is a subset of $\mathbb{R}^m$ is used to determine when two curves γ and η yield the same velocity vector at p. If different curves had different tangent vectors, we could have defined T_pM simply as the collection of all curves passing through p, thus eliminating the need for the ambient space $\mathbb{R}^m$. In any case, it turns out that the equality $\gamma'(0) = \eta'(0)$ can be checked by looking at the two curves γ and η in local coordinates near p, meaning that $\gamma'(0) = \eta'(0)$ if and only if $(\phi\circ\gamma)'(0) = (\phi\circ\eta)'(0)$, where $\phi : U \subset M \to \mathbb{R}^n$ is some coordinate system with $U \ni p$ open. In particular, $\gamma'(0) = \eta'(0)$ if and only if $(f\circ\gamma)'(0) = (f\circ\eta)'(0)$ for all $f : M \to \mathbb{R}$ smooth. This leads to the following intrinsic definition of tangent spaces.

Definition 3.1. Let M be a manifold, and let $W_p(M) = \{\gamma \in \mathscr{C}^\infty(\mathbb{R}, M) : \gamma(0) = p\}$ be the space of curves passing through $p \in M$. Then we define

$$T_pM = W_p(M)/\sim, \tag{20}$$

where the equivalence relation $\sim$ on $W_p(M)$ is defined by

$$\gamma \sim \eta \qquad \Longleftrightarrow \qquad (f\circ\gamma)'(0) = (f\circ\eta)'(0) \quad \text{for all } f \in \mathscr{C}^\infty(M). \tag{21}$$

The elements of T_pM are called the *tangent vectors of M at p*.

Note that if $\gamma, \eta \in W_p(M)$ are curves representing two *different* tangent vectors, then there exists f such that $(f\circ\gamma)'(0) \neq (f\circ\eta)'(0)$. Moreover, we observe that the quantity $(f\circ\gamma)'(0)$ depends only on f and the equivalence class $[\gamma] \in T_pM$, and *not* on the particular representative curve γ. In the embedded case $M \subset \mathbb{R}^m$, we recognize $(f\circ\gamma)'(0)$ as the directional derivative of f along the vector $\gamma'(0)$.

Definition 3.2. For $V \in T_pM$ and $f \in \mathscr{C}^\infty(M)$, we define

$$V(f) = (f\circ\gamma)'(0), \tag{22}$$

where $\gamma \in W_p(M)$ is such that $V = [\gamma]$. This defines a linear map $V : \mathscr{C}^\infty(M) \to \mathbb{R}$, and may be called the *action* of V on f, or the *directional derivative* of f along the direction V.

Example 3.3. Let $\Omega \subset \mathbb{R}^n$ be open, and let $p \in \Omega$. Then for any $f \in \mathscr{C}^\infty(\Omega)$ and $\gamma \in W_p(\Omega)$, we have

$$(f\circ\gamma)'(0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(p)\gamma_i'(0), \tag{23}$$

where $\gamma_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, n$, are the components of $\gamma : \mathbb{R} \to \Omega \subset \mathbb{R}^n$. Thus we see that $(f \circ \gamma)'(0) = (f \circ \eta)'(0)$ for all f if and only if $\gamma'(0) = \eta'(0)$. Moreover, for any $a \in \mathbb{R}^n$, we can construct a curve $\gamma \in W_p(\Omega)$ with $\gamma'(0) = a$. This means that the tangent space $T_p\Omega$ can be identified with $\mathbb{R}^n$ through the correspondence $[\gamma] \mapsto \gamma'(0)$. As for the directional derivative aspect, for $V = [\gamma] \in T_p\Omega$, we get

$$V(f) = (f \circ \gamma)'(0) = \sum_{i=1}^{n} a_i \frac{\partial f}{\partial x_i}(p), \tag{24}$$

where $a = \gamma'(0) \in \mathbb{R}^n$. Since f is arbitrary, we rewrite it as

$$V = \sum_{i=1}^{n} a_i \Big(\frac{\partial}{\partial x_i}\Big)_p, \tag{25}$$

and think of V itself as a differential operator (at a point). Then the partial derivative operators $\frac{\partial}{\partial x_i}$ acting at the point p form a basis of the tangent space $T_p\Omega$.

Remark 3.4. The arguments from the preceding example can be adapted to the general case. Let (U, ϕ) be a coordinate chart on M with $U \ni p$. Then we have

$$V(f) = (f \circ \gamma)'(0) = (f \circ \phi^{-1} \circ \phi \circ \gamma)'(0) = \sum_{i=1}^{n} \frac{\partial (f \circ \phi^{-1})}{\partial x_i}(\phi(p))(\phi_i \circ \gamma)'(0), \tag{26}$$

which shows that the coefficients $(\phi \circ \gamma)'(0) \in \mathbb{R}^n$ completely determine the vector V, in the sense that if $(\phi \circ \gamma)'(0) = (\phi \circ \eta)'(0)$ then $(f \circ \gamma)'(0) = (f \circ \eta)'(0)$ for all f. In other words, the map $\phi_* : T_pM \to \mathbb{R}^n$ defined by $\phi_* V = (\phi \circ \gamma)'(0)$ is injective. On the other hand, given $a \in \mathbb{R}^n$, we can define the curve $\gamma(t) = \phi^{-1}(\phi(p) + ta)$ satisfying $(\phi \circ \gamma)'(0) = a$, meaning that $\phi_* : T_pM \to \mathbb{R}^n$ is surjective. Hence we are justified to call $(\phi \circ \gamma)'(0) \in \mathbb{R}^n$ the *coordinate representation* of $V = [\gamma] \in T_pM$, with respect to the coordinate system ϕ. The invertible map $\phi_* : T_pM \to \mathbb{R}^n$ also induces a linear structure on T_pM, making it an n-dimensional vector space.

Now, if ψ is another coordinate system, then

$$\psi_* V = (\psi \circ \gamma)'(0) = (\psi \circ \phi^{-1} \circ \phi \circ \gamma)'(0) = D(\psi \circ \phi^{-1})(\phi(p))(\phi \circ \gamma)'(0) = J\phi_* V, \tag{27}$$

where J is the Jacobian matrix of $\psi \circ \phi^{-1}$ evaluated at $\phi(p)$. Since this *coordinate transformation law* for vectors is linear, it is clear that the linear structure of T_pM induced by ϕ is identical to the linear structure induced by ψ. With $y = (\psi \circ \phi^{-1})(x)$, we can rewrite the transformation law as

$$b_k = \sum_{i=1}^{n} \frac{\partial y_k}{\partial x_i} a_i, \tag{28}$$

where $a = \phi_* V$ and $b = \psi_* V$.

For any $V \in T_pM$, the directional derivative $V : \mathscr{C}^\infty(M) \to \mathbb{R}$, as in Definition 3.2, can be written as

$$V(f) = \sum_{i=1}^{n} a_i \frac{\partial (f \circ \phi^{-1})}{\partial x_i}(\phi(p)), \tag{29}$$

where $a = \phi_* V$, and hence satisfies the *Leibniz law*

$$V(fg) = f(p)V(g) + g(p)V(f). \tag{30}$$

Such a map $V : \mathscr{C}^\infty(M) \to \mathbb{R}$ is called a *derivation* of $\mathscr{C}^\infty(M)$ at p. It turns out that the derivation property characterizes tangent vectors.

Theorem 3.5. *Let $D : \mathscr{C}^\infty(M) \to \mathbb{R}$ be a derivation of $\mathscr{C}^\infty(M)$ at p, in the sense that D is linear and satisfies the Leibniz law (30). Then D is realized by an element of T_pM.*

Proof. Suppose that $g \in \mathscr{C}^\infty(M)$ vanishes in a neighbourhood of p, and let $h \in \mathscr{C}^\infty(M)$ be such that $h(p) \neq 0$ and $gh \equiv 0$ on M. Then the Leibniz law gives

$$0 = D(gh) = g(p)D(h) + h(p)D(g) = h(p)D(g), \tag{31}$$

that is, $D(g) = 0$. This means that $D(f)$ depends only on the local behaviour of $f \in \mathscr{C}^\infty(M)$ near the point p, and hence we can restrict our attention to a coordinate chart (U, ϕ) on M with $U \ni p$ and $\phi(p) = 0$. Then for $g = f \circ \phi^{-1} \in \mathscr{C}^\infty(U)$, we have

$$g(x) = g(0) + \int_0^1 \frac{d}{dt} g(tx) dt = g(0) + \int_0^1 \sum_{i=1}^n \frac{\partial g}{\partial x_i}(tx) x_i dt = c + \sum_{i=1}^n g_i(x) x_i, \tag{32}$$

where $c = g(0)$, and

$$g_i(x) = \int_0^1 \frac{\partial g}{\partial x_i}(tx) dt, \qquad i = 1, \ldots, n. \tag{33}$$

Note that $g_i(0) = \frac{\partial g}{\partial x_i}(0)$. Now with $f_i = g_i \circ \phi$, we can write

$$f(q) = c + \sum_{i=1}^n f_i(q) \phi_i(q), \qquad q \in U, \tag{34}$$

and by applying D on both sides, we get

$$D(f) = D(c) + \sum_{i=1}^n D(f_i)\phi_i(p) + \sum_{i=1}^n f_i(p) D(\phi_i) = D(c) + \sum_{i=1}^n f_i(p) D(\phi_i), \tag{35}$$

where we have used the fact that $\phi_i(p) = 0$. For the constant function 1, we have

$$D(1) = D(1 \cdot 1) = 1D(1) + 1D(1) = 2D(1), \tag{36}$$

yielding $D(c) = D(c\dot{1}) = cD(1) = 0$ by linearity. Finally, taking into account the equality

$$f_i(p) = g_i(0) = \frac{\partial g}{\partial x_i}(0) = \frac{\partial (f \circ \phi^{-1})}{\partial x_i}(\phi(p)), \tag{37}$$

we infer

$$D(f) = \sum_{i=1}^n f_i(p) D(\phi_i) = \sum_{i=1}^n a_i \frac{\partial (f \circ \phi^{-1})}{\partial x_i}(\phi(p)), \tag{38}$$

where $a_i = D(\phi_i)$. Therefore D is realized by the vector $V = (\phi_*)^{-1} a \in T_p M$. $\square$

Let $\phi : M \to N$ be a smooth map, and let $p \in M$. Then a curve $\gamma \in W_p(M)$ gets sent to the curve $\eta = \phi \circ \gamma \in W_q(N)$, with $q = \phi(p)$. Moreover, for $g \in \mathscr{C}^\infty(N)$, we have

$$(g \circ \eta)'(0) = (g \circ \phi \circ \gamma)'(0) = (f \circ \gamma)'(0), \qquad \text{with} \quad f = g \circ \phi, \tag{39}$$

meaning that ϕ induces a map $\phi_* : T_p M \to T_q N$ between the tangent spaces, by

$$\phi_*[\gamma] = [\phi \circ \gamma]. \tag{40}$$

This is called the *differential of* ϕ *at* p, which may also be denoted by $(\mathrm{d}\phi)_p$, $D\phi(p)$, and $T_p \phi$. From (39), we can further extract an equivalent definition in the language of derivations, as

$$(\phi_* V)(g) = V(g \circ \phi), \tag{41}$$

for $V \in T_p M$ and $g \in \mathscr{C}^\infty(N)$. It makes it clear that

- ϕ_* is a linear map.
- If $\psi : N \to \Sigma$ is another smooth map, then $(\psi \circ \phi)_* = \psi_* \circ \phi_*$.
- In particular, $(\phi^{-1})_* = (\phi_*)^{-1}$ if ϕ is a diffeomorphism.

Remark 3.6. Suppose that $\phi : U \to \Omega$ is a diffeomorphism, where $U \subset M$ and $\Omega \subset \mathbb{R}^n$ are open, i.e., ϕ is a coordinate system. Then for $g \in \mathscr{C}^\infty(\Omega)$ and $\gamma \in W_p(U)$, with $\phi_i : U \to \mathbb{R}$ denoting the i-th component of ϕ, we have

$$(g \circ \phi \circ \gamma)'(0) = \sum_{i=1}^n \frac{\partial g}{\partial x_i}(\phi(p))(\phi_i \circ \gamma)'(0), \tag{42}$$

which yields

$$(\phi_* V)(g) = \sum_{i=1}^n (\phi_i \circ \gamma)'(0) \frac{\partial g}{\partial x_i}(\phi(p)) = \sum_{i=1}^n V(\phi_i) \frac{\partial g}{\partial x_i}(\phi(p)). \tag{43}$$

This can be written as

$$\phi_* V = \sum_{i=1}^n a_i \Big(\frac{\partial}{\partial x_i}\Big)_{\phi(p)} \in T_{\phi(p)}\Omega, \tag{44}$$

where $a_i = V(\phi_i) = (\phi_i \circ \gamma)'(0)$, $i = 1, \ldots, n$, cf. Example 3.3. Since $\phi_* V$ can be identified with $(a_1, \ldots, a_n) \in \mathbb{R}^n$, the differential map $\phi_* : T_p M \to T_{\phi(p)}\Omega$ is consistent with the map $\phi_* : T_p M \to \mathbb{R}^n$ defined in Remark 3.4. Applying $(\phi_*)^{-1}$ on both sides of the preceding equality, we infer

$$V = \sum_{i=1}^n a_i (\phi_*)^{-1} \Big(\frac{\partial}{\partial x_i}\Big)_{\phi(p)}, \tag{45}$$

showing that the vectors $E_i = (\phi_*)^{-1}(\frac{\partial}{\partial x_i})_{\phi(p)} \in T_p M$ form a basis of the tangent space $T_p M$. Note that the vector E_i is represented by the curve $\gamma(t) = \phi^{-1}(\phi(p) + t e_i)$, which is simply the "coordinate line" x_i passing through the point p. For this reason, it is convenient to introduce the notation $(\frac{\partial}{\partial x_i})_p \equiv E_i$, and rewrite the preceding equation as

$$V = \sum_{i=1}^n a_i \Big(\frac{\partial}{\partial x_i}\Big)_p. \tag{46}$$

The collection $\{(\frac{\partial}{\partial x_i})_p : i = 1, \ldots, n\}$ is called the *coordinate (or holonomic) basis* of $T_p M$, associated to the coordinate system ϕ.

Example 3.7. Let E be a vector space, and let $p \in E$. Let $\tau : E \to E$ be defined by $\tau(x) = x - p$. Clearly, we have $\tau(p) = 0$ and $\tau^{-1}(x) = x + p$, which ensures that $\tau_* : T_p E \to T_0 E$ is an isomorphism. Thus $T_p E$ is naturally identified with $T_0 E$. We shall further identify $T_0 E$ with E itself. Let $\phi : E \to \mathbb{R}^n$ be an isomorphism. Then the vectors $(\phi_*)^{-1}(\frac{\partial}{\partial x_i})_0 \in T_0 E$ form a basis of $T_0 E$. In other words, $\phi_* : T_0 E \to T_0 \mathbb{R}^n$ is a linear isomorphism. Since $T_0 \mathbb{R}^n$ can be canonically identified with $\mathbb{R}^n$, we can write $\phi_* : T_0 E \to \mathbb{R}^n$, inducing the identification $\phi^{-1} \circ \phi_* : T_0 E \to E$. We claim that this identification does not depend on ϕ. Let $\psi : E \to \mathbb{R}^n$ be another isomorphism. Then we have $\psi_* = (\psi \circ \phi^{-1})_* \circ \phi_* = (\psi \circ \phi^{-1}) \circ \phi_*$, because $(\psi \circ \phi^{-1})_* = \psi \circ \phi^{-1}$ under the identification between $T_0 \mathbb{R}^n$ and $\mathbb{R}^n$. Hence we infer $\psi^{-1} \circ \psi_* = \psi^{-1} \circ (\psi \circ \phi^{-1}) \circ \phi_* = \phi^{-1} \circ \phi_*$, which proves the claim.

Exercise 3.1. Let $\phi : M \to N$ be a smooth map, and let $p \in M$. Find the expression for $(\mathrm{d}\phi)_p$ in a coordinate basis.

Exercise 3.2. Show that the *tangent bundle* $TM = \{(p, v) : p \in M,\, v \in T_p M\}$ admits a canonical smooth structure, making it a smooth manifold.

4. Vector fields and flows

A vector at $p \in M$ is simply the operation of taking the derivative $(f \circ \gamma)'(0)$ of a scalar function f, for some curve $\gamma \in W_p(M)$. Moreover, for any smooth curve $\gamma : (a, b) \to M$ and any parameter value $t \in (a, b)$, we can define the vector $\gamma'(t) \in T_{\gamma(t)}M$ by

$$\gamma'(t)f = (f \circ \gamma)'(t) = \Big(\frac{d}{dt}\Big)_t (f \circ \gamma), \tag{47}$$

where $(\frac{d}{dt})_t$ is the derivative operator acting at the point $t \in (a, b)$, in the same spirit as the notation $(\frac{\partial}{\partial x_i})_p$. We call $\gamma'(t)$ the *velocity vector* of γ at the parameter value t. Since $(\frac{d}{dt})_t \in T_t(a, b)$, we can also write

$$\gamma'(t) = \gamma_* \Big(\frac{d}{dt}\Big)_t = (\mathrm{d}\gamma)_t \Big(\frac{d}{dt}\Big)_t. \tag{48}$$

If we want to consider the problem of finding a curve with given velocity vectors, or if we want to be able to take a directional derivative of scalar functions at each point of M, then we need to consider a *vector field* X, which is by definition an assignment of a vector $X_p \in T_pM$ to each point $p \in M$. Unless otherwise specified, vector fields are assumed to be *smooth*, in the sense that $Xf \in \mathscr{C}^\infty(M)$ for all $f \in \mathscr{C}^\infty(M)$, where the function Xf is defined by pointwise application of X to f, as

$$(Xf)(p) = X_p(f), \qquad p \in M. \tag{49}$$

The set of all smooth vector fields on M is denoted by $\mathfrak{X}(M)$ or $\mathscr{C}^\infty(TM)$.

In local coordinates $\phi : U \subset M \to \mathbb{R}$, we have

$$(Xf)(p) = \sum_{i=1}^n a_i(p) \frac{\partial f}{\partial x_i}(p), \tag{50}$$

with $a_i(p) = X_p(\phi_i)$, and since the components ϕ_i of ϕ are smooth, the coefficients a_i of X are smooth functions in U. On the other hand, if all coefficients a_i are smooth, then from the preceding formula we see that Xf is smooth in U, for all smooth f. To conclude, smoothness of a vector field is the same as smoothness of its coefficients in local coordinates.

Remark 4.1. In $\mathfrak{X}(M)$, addition, subtraction, and multiplication by scalar functions are well-defined, meaning that $\mathfrak{X}(M)$ is a module over $\mathscr{C}^\infty(M)$. Moreover, any $X \in \mathfrak{X}(M)$ is a *derivation of* $\mathscr{C}^\infty(M)$, in the sense that $X : \mathscr{C}^\infty(M) \to \mathscr{C}^\infty(M)$ is linear, and satisfies the Leibniz law

$$X(fg) = fXg + gXf, \qquad f, g \in \mathscr{C}^\infty(M). \tag{51}$$

The derivation property characterizes vector fields. To see this, let $D : \mathscr{C}^\infty(M) \to \mathscr{C}^\infty(M)$ be a derivation. Then at $p \in M$, it satisfies

$$D(fg)|_p = f(p)(Dg)|_p + g(p)(Df)|_p, \tag{52}$$

and so by Theorem 3.5, there is $X_p \in T_pM$ such that $D(f)|_p = X_p(f)$ for all f. This defines a vector field $X \in \mathfrak{X}(M)$, which is smooth, since $Xf = D(f)$ is smooth for all $f \in \mathscr{C}^\infty(M)$.

Now we start our discussion about finding curves with given velocity vectors.

Definition 4.2. A curve $\gamma : (a, b) \to M$ is called an *integral curve* of $X \in \mathfrak{X}(M)$, if

$$\gamma'(t) = X_{\gamma(t)} \qquad \text{for all} \quad t \in (a, b). \tag{53}$$

It is called *maximal* (or inextendible), if there is no integral curve $\eta : (\alpha, \beta) \to M$ of X satisfying $\gamma = \eta|_{(a,b)}$ and $(a, b) \subsetneq (\alpha, \beta)$.

To be concrete, let $X \in \mathfrak{X}(M)$, let $p \in M$, and consider the problem

$$\begin{cases} \gamma'(t) = X_{\gamma(t)} & \text{for } t \in (a, b), \\ \gamma(0) = p. \end{cases} \tag{54}$$

In local coordinates $\phi : U \to \Omega$, where $U \subset M$ and $\Omega \subset \mathbb{R}^n$ are open, if we write $F = \phi_* X$ and $x(t) = \phi(\gamma(t)) \in \Omega$, this problem is equivalent to the initial value problem

$$\begin{cases} x'(t) = F(x(t)) & \text{for } t \in (a, b), \\ x(0) = q, \end{cases} \tag{55}$$

where $q = \phi(p) \in \Omega$. Then the standard ODE theory tells us the following.

- There exists a maximal solution $x : (a, b) \to \Omega$ of (55), with $a < 0 < b$.
- This maximal solution is unique.
- For each compact set $K \subset \Omega$, there exists $\varepsilon > 0$, such that $(a, b) \supset (-\varepsilon, \varepsilon)$ independent of $q \in K$. In particular, if $b < \infty$ (respectively, if $a > -\infty$), then $x(t)$ escapes every compact set $K \subset \Omega$ as $t \nearrow b$ (respectively, as $t \searrow a$).
- Let us write (a_q, b_q) and $x_q(t)$ to make the dependence of these quantities on the initial data q explicit. Then the set $\Sigma = \{(q, t) : q \in \Omega,\ a_q < t < b_q\}$ is open, and the map $(q, t) \mapsto x_q(t)$ is smooth in Σ.

It is not difficult to generalize the aforementioned results to the problem (54).

Theorem 4.3. *Given any manifold M, a vector field $X \in \mathfrak{X}(M)$, and a point $p \in M$, there exists a unique maximal integral curve $\gamma : (a, b) \to M$ satisfying (54), with $a < 0 < b$. Furthermore, we have the following.*

(a) For each compact set $K \subset M$, there exists $\varepsilon > 0$, such that $(a, b) \supset (-\varepsilon, \varepsilon)$ independent of $p \in K$. In particular, if $b < \infty$ (respectively, if $a > -\infty$), then $\gamma(t)$ escapes every compact set $K \subset M$ as $t \nearrow b$ (respectively, as $t \searrow a$).

(b) Let us write (a_p, b_p) and $\gamma_p(t)$ to make the dependence of these quantities on the initial point p explicit. Then the set $\Sigma = \{(p, t) : p \in M,\ a_p < t < b_p\}$ is open, and the map $(p, t) \mapsto \gamma_p(t)$ is smooth as a map between Σ and M.

Proof. Fix $p \in M$, and let $\{I_\alpha\}$ be the set of all intervals such that there is a solution $\gamma_\alpha : I_\alpha \to M$ of the initial value problem (54), where α runs over some index set. Consider two indices α and β, and let $t \in I_\alpha \cap I_\beta$. Without loss of generality, assume $t > 0$, and cover the compact set $\gamma_\alpha([0, t])$ by a finitely many coordinate charts $U_1, \ldots, U_m$, with $p \in U_1$ and $\gamma_\alpha(t) \in U_m$. Since $\gamma_\alpha(0) = \gamma_\beta(0)$, by local theory, we have $\gamma_\alpha \equiv \gamma_\beta$ in U_1. Then, assuming that $U_1 \cap U_2 \cap \gamma_\alpha([0, t]) \neq \varnothing$, the same argument implies $\gamma_\alpha \equiv \gamma_\beta$ in U_2. Repeating this, we get $\gamma_\alpha(t) = \gamma_\beta(t)$. Thus, we can consistently define a function $\gamma : I \to M$ with $I = \bigcup_\alpha I_\alpha$, satisfying $\gamma(t) = \gamma_\alpha(t)$ for all α, whenever $t \in I_\alpha$. By construction, γ is a solution of the initial value problem (54), and it cannot be extended to a larger interval of existence, hence its maximality. For uniqueness, suppose that $\gamma_1 : I_1 \to M$ and $\gamma_2 : I_2 \to M$ are two maximal solutions. Then we can construct an extension γ defined over $I_1 \cup I_2$. However, by maximality of γ_1, we get $I_1 = I$, and similarly, $I_2 = I$. This gives $I_1 = I_2$, and hence $\gamma_1 = \gamma_2$.

To prove (b), let $(p, t) \in \Sigma$. Without loss of generality, assume that $t > 0$, and let $U_1, \ldots, U_m$ be coordinate charts covering $\gamma_p([0, t])$ as in the preceding paragraph. Then we subdivide the time interval $[0, t]$ into smaller subintervals $[0, t_1]$, $[t_1, t_2]$, and so on, until $[t_N, t]$, with the image of each subinterval under γ_p entirely contained in one of the charts $U_1, \ldots, U_m$. Let $p_i = \gamma_p(t_i)$ for all i. First, we choose an open set $V_N \subset M$ containing p_N, such that $\gamma_q(t - t_N)$ is defined for all $q \in V_N$. This is possible by the ODE theory in $\mathbb{R}^n$. In the second step, we choose an open set $V_{N-1} \subset M$ containing p_{N-1}, such that $\gamma_q(t_N - t_{N-1})$ is not only defined, but also satisfies $\gamma_q(t_N - t_{N-1}) \in V_N$ for all $q \in V_{N-1}$. We continue this process, until we reach the point p, where we choose an open set $V_0 \subset M$ containing p, such that $\gamma_q(t_1)$ is defined and $\gamma_q(t_1) \in V_1$ for all $q \in V_0$. Now, if we write $\Phi_\tau(q) = \gamma_q(\tau)$, then by construction, $\Phi_t : V_0 \to M$ is well defined, as it is the composition of finitely many smooth maps. Finally, let $\Sigma_0 \subset M \times \mathbb{R}$ be an open set containing $(p, 0)$, such that $\gamma_q(\tau)$ is defined and $\gamma_q(\tau) \in V_0$ for all $(q, \tau) \in \Sigma_0$. Then $\Sigma_t = \{(q, \tau + t) : (q, \tau) \in \Sigma_0\}$ is an open neighbourhood of (p, t), and for any $(q, \tau) \in \Sigma_t$, we have $\gamma_q(\tau) = \Phi_t(\gamma_q(\tau - t))$, which makes it clear that $\gamma_q(\tau)$ is well defined, and is a smooth function of $(q, \tau) \in \Sigma_t$.

As for (a), consider a compact set $K \subset M$. Invoking Theorem 2.13, we can find a *finite* collection of diffeomorphisms $\psi_i : B_1 \to V_i \subset M$, such that $\{\psi_i(B_{1/2})\}$ is a cover of K. Since $\overline{B}_{1/2} \subset B_1$ is compact, the claim follows from the local results. □

A vector field is called *complete* if each of its integral curves is defined for all $t \in \mathbb{R}$.

Corollary 4.4. *Any vector field on a compact manifold is complete.*

Example 4.5. Let $M = \mathbb{R}$, and consider the initial value problems $x' = x$ and $x' = x^2$, with $x(0) = q$. The solution of the first problem is $x(t) = qe^t$. Obviously, we have in this case $(a, b) = \mathbb{R}$ for all $q \in \mathbb{R}$, meaning that the vector field $X = x\frac{\partial}{\partial x}$ is complete. The second problem is solved by $x(t) = \frac{q}{1-qt}$, which *blows up* as $t \nearrow b = \frac{1}{q}$ for $q > 0$ and as $t \searrow a = \frac{1}{q}$ for $q < 0$. Thus, the maximal interval of existence is given by

$$(a, b) = \begin{cases} \mathbb{R} & \text{for } q = 0, \\ (-\infty, \frac{1}{q}) & \text{for } q > 0, \\ (\frac{1}{q}, \infty) & \text{for } q < 0, \end{cases} \tag{56}$$

which shows that the vector field $X = x^2\frac{\partial}{\partial x}$ is incomplete. More generally, $x^{1+\alpha}\frac{\partial}{\partial x}$ is incomplete as long as $\alpha > 0$. On the other hand, the vector fields $x \log x \frac{\partial}{\partial x}$ and $x(\log x) \log \log x \frac{\partial}{\partial x}$ on $\mathbb{R}$ are complete, even though their integral curves behave like e^{e^t} and $e^{e^{e^t}}$, respectively. One can add more iterated logarithms without making the vector field incomplete, but any of those logarithms cannot be raised to a power greater than 1. For example, $x(\log x)^{1+\alpha}\frac{\partial}{\partial x}$ and $x(\log x)(\log \log x)^{1+\alpha}\frac{\partial}{\partial x}$ are incomplete, as long as $\alpha > 0$.

Definition 4.6. Let M be a manifold, and let $X \in \mathfrak{X}(M)$. For $p \in M$, let $\gamma_p : (a_p, b_p) \to M$ be the maximal integral curve satisfying (54). Then the *flow (map) of* X is defined by $\Phi_t(p) = \gamma_p(t)$ for $p \in M$ and $a_p < t < b_p$.

Since $\gamma_p(0) = p$, we have $\Phi_0 = \mathrm{id}$ on M. Given $t \in \mathbb{R}$, the flow map Φ_t is defined on $\Sigma_t = \{p \in M : t \in (a_p, b_p)\}$, which is in general a proper subset of M, but is nonempty for small t. We also know from Theorem 4.3 that the set $\Sigma = \{(p, t) : p \in M, a_p < t < b_p\}$ is open, and the map $(p, t) \mapsto \Phi_t(p)$ is smooth as a map between Σ and M.

Theorem 4.7. *For any $p \in M$, there exist an open set $p \in U \subset M$ and $\varepsilon > 0$, such that Φ_t is defined on U for all $t \in (-\varepsilon, \varepsilon)$, and*

$$\Phi_{s+t} = \Phi_t \circ \Phi_s \qquad \text{on} \quad U, \tag{57}$$

whenever $s, s + t \in (-\varepsilon, \varepsilon)$. In particular, for $t \in (-\varepsilon, \varepsilon)$, the flow map $\Phi_t : U \to \Phi_t(U)$ is a diffeomorphism with $(\Phi_t)^{-1} = \Phi_{-t}$.

Proof. Let $p \in M$ and $s \in (a_p, b_p)$. Define $\eta(t) = \gamma_p(s + t)$ for $s + t \in (a_p, b_p)$. Then we have

$$\eta'(t) = \gamma_p'(s + t) = X_{\gamma_p(s+t)} = X_{\eta(t)}, \qquad \eta(0) = \gamma_p(s), \tag{58}$$

which shows that $t \in (a_q, b_q)$ and $\eta(t) = \gamma_q(t)$, where $q = \gamma_p(s)$. In other words, we have

$$\Phi_{s+t}(p) = \Phi_t(\Phi_s(p)) \qquad \text{whenever} \quad s, s + t \in (a_p, b_p). \tag{59}$$

Given $p \in M$, we can find an open set $p \in U \subset M$ and $\varepsilon > 0$, such that $U \times (\varepsilon, \varepsilon) \subset \Sigma$, where Σ is the domain of the map $(q,t) \mapsto \Phi_t(q)$, cf. Theorem 4.3 or the paragraph preceding the current theorem. This completes the proof. $\square$

Exercise 4.1. Let $X \in \mathfrak{X}(M)$, and let Φ_t be the associated flow. Show that for $\lambda \in \mathbb{R} \setminus \{0\}$, the flow associated to the scaled vector field λX is given by $\Phi_{t/\lambda}$.

5. Lie derivatives

Recall that the differential of a smooth map $\phi : M \to N$ at $p \in M$ is a map between the tangent spaces T_pM and T_qN, which is defined by

$$(\phi_* V)(g) = V(g \circ \phi) \qquad \text{for} \quad V \in T_pM \quad \text{and} \quad g \in \mathscr{C}^\infty(N), \tag{60}$$

or equivalently, by

$$\phi_*[\gamma] = [\phi \circ \gamma] \qquad \text{for} \quad \gamma \in W_p(M). \tag{61}$$

By allowing $p \in M$ to vary, we can regard ϕ_* as a mapping between the tangent bundles TM and TN. However, in general, this map does not send a vector field X on M into a vector field Y on N, since (i) ϕ may be nonsurjective, in which case we have no way of unambiguously defining Y everywhere on N, and (ii) ϕ may be noninjective, meaning that there exist distinct points $p, q \in M$ satisfying $\phi(p) = \phi(q)$, and so $\phi_* X_p \neq \phi_* X_q$ in general. For invertible mappings, all these problems disappear.

Definition 5.1. If $\phi : M \to N$ is a diffeomorphism between two manifolds, then the *push-forward* $\phi_* X \in \mathfrak{X}(N)$ of $X \in \mathfrak{X}(M)$ is defined by $(\phi_* X)_{\phi(p)} = (\mathrm{d}\phi)_p X_p$, that is,

$$(\phi_* X) f = X(f \circ \phi) \circ \phi^{-1} \qquad \text{for} \quad f \in \mathscr{C}^\infty(N). \tag{62}$$

Let us compute the push-forward in local coordinates. Thus let $p \in M$, and let $\{x_i\}$ and $\{y_k\}$ be coordinate systems on neighbourhoods of p and of $q = \phi(p)$, respectively. With X_i and ϕ_k denoting the components of X and ϕ, respectively, we have

$$(\phi_* X)_q f = X_p(f \circ \phi) = \sum_i (X_i)_p \Big(\frac{\partial \phi_k}{\partial x_i}\Big)_p \Big(\frac{\partial f}{\partial y_k}\Big)_q, \tag{63}$$

yielding

$$(\phi_* X)_q = \sum_i (X_i)_p \Big(\frac{\partial \phi_k}{\partial x_i}\Big)_p \Big(\frac{\partial}{\partial y_k}\Big)_q. \tag{64}$$

Let $X \in \mathfrak{X}(M)$ and let Φ_t be the flow of X. Fix a point $p \in M$. Then by Theorem 4.7, there exist an open set $U \ni p$ and $\varepsilon > 0$ such that $\Phi_t : U \to \Phi_t(U)$ is a diffeomorphism with $(\Phi_t)^{-1} = \Phi_{-t}$ for each $t \in (-\varepsilon, \varepsilon)$. Now let $Y \in \mathfrak{X}(M)$ be another vector field. Then $(\Phi_{-t})_*(Y|_{\Phi_t(U)})$ would be a vector field on U, that, in a certain sense, can be considered as the vector field $Y|_{\Phi_t(U)}$ "transported back" to U along the vector field X. In particular, by writing $(\Phi_{-t})_* Y = (\Phi_{-t})_*(Y|_{\Phi_t(U)})$ for simplicity, at the point p, we can think of the vector $\Gamma(t) = ((\Phi_{-t})_* Y)_p \in T_pM$ as the vector $Y_{\Phi_t(p)}$ brought back to p along the vector field X. In local coordinates, we have

$$\Gamma(t) = \sum_i (Y_i)_{\Phi_t(p)} \Big(\frac{\partial (\Phi_{-t})_k}{\partial x_i}\Big)_{\Phi_t(p)} \Big(\frac{\partial}{\partial x_k}\Big)_p, \tag{65}$$

which shows that $\Gamma : (-\varepsilon, \varepsilon) \to T_pM$ is a smooth curve. Hence $\Gamma'(0)$ is well defined, and is an element of the tangent space of T_pM at Y_p. Since $T_{Y_p}T_pM$ can be canonically identified with T_pM, we can (and will) consider $\Gamma'(0)$ as an element of T_pM.

Definition 5.2. In the setting of the preceding paragraph, the *Lie derivative of* $Y \in \mathfrak{X}(M)$ *along* $X \in \mathfrak{X}(M)$ is defined by

$$\mathfrak{L}_X Y = \frac{d}{dt}((\Phi_{-t})_* Y)\Big|_{t=0}. \tag{66}$$

Remark 5.3. For scalar functions, we may introduce the Lie derivative by using the notion of *pull-back* $\Phi_t^* f = f \circ \Phi_t$ as

$$\mathfrak{L}_X f = \frac{d}{dt}(\Phi_t^* f)\Big|_{t=0}. \tag{67}$$

However, since $\Phi_t^* f(p) = f(\Phi_t(p)) = f(\gamma_p(t))$, we get

$$(\mathfrak{L}_X f)(p) = \frac{d}{dt} f(\gamma_p(t))\Big|_{t=0} = (f \circ \gamma_p)'(0) = X_p f, \tag{68}$$

meaning that the Lie derivative for scalar functions is simply the directional derivative.

Let us get back to the Lie derivatives of vector fields. We have

$$\begin{aligned}((\Phi_{-t})_* Y)f - Yf &= Y(f \circ \Phi_{-t}) \circ \Phi_t - Y(f) \circ \Phi_t + Y(f) \circ \Phi_t - Yf \\ &= Y(\Phi_{-t}^* f - f) \circ \Phi_t + \Phi_t^*(Yf) - Yf,\end{aligned} \tag{69}$$

and dividing through by t and taking the limit $t \to 0$, we get

$$\mathfrak{L}_X Y = Y(-Xf) + X(Yf) = X(Yf) - Y(Xf). \tag{70}$$

Definition 5.4. The *Lie bracket* or *commutator* of $X, Y \in \mathfrak{X}(M)$ is a vector field defined by

$$[X, Y](f) = X(Yf) - Y(Xf) \qquad \text{for} \quad f \in \mathscr{C}^\infty(M). \tag{71}$$

Clearly, $[X, Y]$ is bilinear and antisymmetric in X and Y. The fact

$$\mathfrak{L}_X Y = [X, Y], \tag{72}$$

already shows that $[X, Y] \in \mathfrak{X}(M)$, but we can independently check if it obeys the Leibniz law as follows. We have

$$X(Y(fg)) = X(fYg + gYf) = fX(Yg) + (Xf)(Yg) + (Xg)(Yf) + gX(Yf), \tag{73}$$

and similarly for $Y(X(fg))$, yielding the desired identity

$$[X, Y](fg) = f[X, Y](g) + g[X, Y](f). \tag{74}$$

Another important property is the following *derivation property*

$$\mathfrak{L}_X(fY) = [X, fY] = X(f)Y + f[X, Y] = X(f)Y + f\mathfrak{L}_X Y, \tag{75}$$

which can be verified direclty.

Example 5.5. Let $\phi : U \to \mathbb{R}^n$ be a coordinate system in an open set $U \subset M$, and let $E_i = (\phi_*)^{-1}(\frac{\partial}{\partial x_i})$, $i = 1, \ldots, n$, be the associated basis of $T_p M$, as $p \in U$ varies. Then for $f \in \mathscr{C}^\infty(M)$, we have

$$E_i f = \frac{\partial (f \circ \phi^{-1})}{\partial x_i}, \qquad \text{and} \qquad E_j E_i f = \frac{\partial^2 (f \circ \phi^{-1})}{\partial x_i \partial x_j}, \tag{76}$$

which implies that

$$[E_i, E_j] = 0 \qquad \text{or simply} \qquad [\tfrac{\partial}{\partial x_i}, \tfrac{\partial}{\partial x_j}] = 0 \qquad \text{for all} \quad i, j. \tag{77}$$

Now let

$$X = \sum_i X_i E_i \in \mathfrak{X}(U), \qquad \text{and} \qquad Y = \sum_i Y_i E_i \in \mathfrak{X}(U). \tag{78}$$

Then we have

$$[X, Y] = \sum_k X(Y_k)E_k + \sum_k Y_k[X, E_k] = \sum_k X(Y_k)E_k - \sum_i Y(X_i)E_i, \tag{79}$$

which gives the coordinate expression of the commutator.

Example 5.6. In $\mathbb{R}^2$, take the two vector fields $X = \frac{\partial}{\partial x}$ and $Y = ax\frac{\partial}{\partial y}$, where $a \in \mathbb{R}$ is a constant, and xy denotes the standard coordinate system in $\mathbb{R}^2$. We have

$$\mathfrak{L}_X Y = -\mathfrak{L}_Y X = [X, Y] = a\frac{\partial}{\partial y}. \tag{80}$$

Note that on the y-axis we have $Y \equiv 0$, but $\mathfrak{L}_Y X \neq 0$ if $a \neq 0$. Thus the Lie derivative $\mathfrak{L}_Y X$ along the y-axis depends on the behaviour of Y "nearby" the y-axis.

Remark 5.7. Let X and Y be vector fields, and let γ be an integral curve of X with $\gamma(0) = p \in M$. Suppose that $Y = 0$ along γ, i.e., $Y_{\gamma(t)} = 0$ for all t. Then for any scalar function f, we have $Yf = 0$ along γ, and hence

$$[X, Y]_p(f) = X_p(Yf) - Y_p(Xf) = ((Yf) \circ \gamma)'(0) = 0. \tag{81}$$

This means that in general, $[X, Y]_p$ depends only on the values of Y *restricted to the curve* γ, or, put it differently, if Y is defined only along the curve γ, then $[X, Y]_p$ does not depend on how one extends Y to be a vector field on M. Note that this behaviour is already apparent from the definition of the Lie derivative.

The following result characterizes commuting vector fields.

Theorem 5.8. *Let X and Y be vector fields on M, and let Φ_t and Ψ_s be the respective flows. Then the following are equivalent.*
(a) $[X, Y] = 0$, that is, the vector fields commute.
(b) $(\Phi_t)_* Y = Y$, that is, Y is invariant along the flow of X.
(c) $\Phi_t \circ \Psi_s = \Psi_s \circ \Phi_t$, that is, the flows commute.

Proof. It is immediate from the definition of the Lie derivative that (b) implies (a). Suppose that (c) holds, and for $p \in M$, write

$$((\Phi_t)_* Y)_p = [\Phi_t \circ \eta_{\Phi_{-t}(p)}], \tag{82}$$

where η_q is the integral curve of Y with $\eta_q(0) = q \in M$, and the equivalence class is understood to be of curves passing through p. By using the commutativity of the flows, we infer

$$\Phi_t(\eta_{\Phi_{-t}(p)}(s)) = \Phi_t(\Psi_s(\Phi_{-t}(p))) = \Phi_t(\Phi_{-t}(\Psi_s(p))) = \Psi_s(p) = \eta_p(s), \tag{83}$$

which proves (b), as

$$((\Phi_t)_* Y)_p = [\Phi_t \circ \eta_{\Phi_{-t}(p)}] = [\eta_p] = Y_p. \tag{84}$$

Finally, suppose that (a) holds. First, with $\Theta_s = \Phi_t \circ \Psi_s \circ \Phi_{-t}$, we have

$$\frac{d}{ds}\Theta_s(p) = (\Phi_t)_*(Y_{\Phi_{-t}(p)}) = ((\Phi_t)_* Y)_p, \tag{85}$$

meaning that Θ_s is the flow of $(\Phi_t)_* Y$. On the other hand, since $(\Phi_{t+\tau})_* = (\Phi_\tau)_*(\Phi_t)_*$ by the group property of the flow, we have

$$\frac{d}{dt}(\Phi_t)_* Y = \frac{d}{d\tau}((\Phi_{t+\tau})_* Y)\Big|_{\tau=0} = -\mathfrak{L}_X(\Phi_t)_* Y = [(\Phi_t)_* Y, X]. \tag{86}$$

Taking into account that

$$(Xf \circ \Phi_t)(p) = X_{\gamma_p(t)} f = (f \circ \gamma_p)'(t) = X(f \circ \Phi_t)(p), \tag{87}$$

the commutator can be computed as

$$\begin{aligned}[(\Phi_t)_*Y, X](f) &= Y(Xf \circ \Phi_t) \circ \Phi_{-t} - X(Y(f \circ \Phi_t) \circ \Phi_{-t}) \\ &= Y(X(f \circ \Phi_t)) \circ \Phi_{-t} - X(Y(f \circ \Phi_t)) \circ \Phi_{-t} \\ &= (\Phi_t)_*[Y, X](f),\end{aligned} \tag{88}$$

and therefore we infer

$$\frac{d}{dt}(\Phi_t)_*Y = (\Phi_t)_*[Y, X] = 0. \tag{89}$$

This shows that $(\Phi_t)_*Y = Y$, and hence $\Phi_t \circ \Psi_s \circ \Phi_{-t} = \Psi_s$, or $\Phi_t \circ \Psi_s = \Psi_s \circ \Phi_t$. □

Exercise 5.1. Prove the *Jacobi identity*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0. \tag{90}$$

Exercise 5.2. Show that

$$\mathfrak{L}_{[X,Y]} = [\mathfrak{L}_X, \mathfrak{L}_Y] := \mathfrak{L}_X \circ \mathfrak{L}_Y - \mathfrak{L}_Y \circ \mathfrak{L}_X. \tag{91}$$

Exercise 5.3. Let X and Y be vector fields on M, and let Φ_t and Ψ_s be the respective flows. Fix a point $p \in M$, and let $\eta(t) = (\Psi_{-t} \circ \Phi_{-t} \circ \Psi_t \circ \Phi_t)(p)$ for t small. Show that

- $\eta'(0) = 0$.
- $(f \circ \eta)''(0) = 2[X, Y]_p(f)$ for any scalar function f.

6. Covectors and 1-forms

The notion of tangent spaces is connected to curves, or maps of the form $\gamma : \mathbb{R} \to M$. At each point $s \in \mathbb{R}$, and under the canonical identification $T_s\mathbb{R} = \mathbb{R}$, the differential of γ is a linear map $(\mathrm{d}\gamma)_s : \mathbb{R} \to T_pM$, where $p = \gamma(s)$. By linearity, this map is completely determined by its value at $1 \in \mathbb{R}$, and we call $\gamma'(s) = (\mathrm{d}\gamma)_s 1$ the velocity vector of γ at s.

An equally natural class is scalar functions $f : M \to \mathbb{R}$, which plays a dual role to curves. The derivative of such a function at $p \in M$ is a linear map $(\mathrm{d}f)_p : T_pM \to \mathbb{R}$, that is, $(\mathrm{d}f)_p$ is an element of the *dual space* of T_p. We say that $(\mathrm{d}f)_p$ is a *covector at* p. The dual space of T_p (or the space of covectors at p) is called the *cotangent space of* M *at* p, and denoted by

$$T_p^*M = (T_pM)^*. \tag{92}$$

Let us look closely at what exactly $(\mathrm{d}f)_p$ is. Before invoking the identification $T_s\mathbb{R} = \mathbb{R}$, the differential is a map $(\mathrm{d}f)_p : T_pM \to T_s\mathbb{R}$ with $s = f(p)$, and by definition, we have

$$((\mathrm{d}f)_pV)(g) = V(g \circ f), \tag{93}$$

for $V \in T_pM$ and $g \in \mathscr{C}^\infty(\mathbb{R})$. In local coordinates, we compute

$$((\mathrm{d}f)_pV)(g) = V(g \circ f) = \sum_i V_i \frac{\partial(g \circ f)}{\partial x_i}\Big|_p = g'(s) \sum_i V_i \frac{\partial f}{\partial x_i}\Big|_p = V(f)\big(\frac{d}{ds}\big)_s(g), \tag{94}$$

and therefore

$$(\mathrm{d}f)_pV = V(f). \tag{95}$$

As an immediate application, for $f, g \in \mathscr{C}^\infty(M)$, we have

$$(\mathrm{d}(fg))_pV = V(fg) = f(p)V(g) + g(p)V(f) = f(p)(\mathrm{d}g)_pV + g(p)(\mathrm{d}f)_pV, \tag{96}$$

meaning that

$$\mathrm{d}(fg) = f\mathrm{d}g + g\mathrm{d}f. \tag{97}$$

Example 6.1. Let $x_i : U \to \mathbb{R}$, $i = 1, \ldots, n$, be coordinate functions with $U \subset M$ open. Since $(\mathrm{d}x_i)_p V = V(x_i)$ for any $V \in T_pM$, we have

$$(\mathrm{d}x_i)_p\Big(\frac{\partial}{\partial x_k}\Big)_p = \Big(\frac{\partial}{\partial x_k}\Big)_p x_i = \Big(\frac{\partial x_i}{\partial x_k}\Big)_p = \delta_{ik}, \tag{98}$$

meaning that $\{(\mathrm{d}x_i)_p\}$ is a basis of T_p^*M, dual to the basis $\{(\frac{\partial}{\partial x_k})_p\}$ of T_pM. We say the the basis $\{(\mathrm{d}x_i)_p\}$ is *induced* by the coordinate system $\{x_i\}$.

Just as vectors lead to the notion of vector fields, covectors lead to the notion of "covector fields," which is customarily called 1-*forms*. More precisely, a 1-form is an assignment of a covector $\alpha_p \in T_p^*M$ to each point $p \in M$. Unless otherwise specified, 1-forms are assumed to be *smooth*, in the sense that $\alpha X \in \mathscr{C}^\infty(M)$ for all $X \in \mathfrak{X}(M)$, where the scalar function αX is defined by pointwise application of α to X, as

$$(\alpha X)(p) = \alpha_p X_p, \qquad p \in M. \tag{99}$$

We will also employ the notations

$$\langle \alpha, X\rangle = \alpha(X) = \alpha X, \qquad \text{and} \qquad \langle \alpha, X\rangle_p = \alpha(X)_p = (\alpha X)(p) = \alpha_p X_p. \tag{100}$$

The set of all smooth 1-forms on M is denoted by $\Omega^1(M)$ or $\mathscr{C}^\infty(T^*M)$.

In local coordinates $x_i : U \subset M \to \mathbb{R}$, $i = 1, \ldots, n$, we have

$$\alpha_p = \sum_{i=1}^n a_i(p)(\mathrm{d}x_i)_p, \tag{101}$$

and applying both sides to $(\frac{\partial}{\partial x_k})_p$ yields $a_i(p) = \alpha_p(\frac{\partial}{\partial x_i})_p$. Since $\frac{\partial}{\partial x_i} \in \mathfrak{X}(U)$, the coefficients a_i of α are smooth functions in U. On the other hand, if all coefficients a_i are smooth, then from the preceding formula we see that αX is smooth in U, for all $X \in \mathfrak{X}(M)$. Therefore, smoothness of a 1-form is the same as smoothness of its coefficients in local coordinates.

In $\Omega^1(M)$, addition, subtraction, and multiplication by scalar functions are well-defined, meaning that $\Omega^1(M)$ is a module over $\mathscr{C}^\infty(M)$. Furthermore, for any 1-form α, the map $X \mapsto \alpha(X) : \mathfrak{X}(M) \to \mathscr{C}^\infty(M)$ is linear and satisfies $\alpha(fX) = f\alpha(X)$ for all $f \in \mathscr{C}^\infty(M)$ and $X \in \mathfrak{X}(M)$. This is called the *tensorial* (or $\mathscr{C}^\infty$*-linearity*) property, which runs a sort of parallel to the derivation property of vector fields, cf. Remark 4.1.

Theorem 6.2. *Let $A : \mathfrak{X}(M) \to \mathscr{C}^\infty(M)$ be a tensorial (or $\mathscr{C}^\infty$-linear) map, in the sense that A is linear and satisfies*

$$A(fX) = fA(X), \qquad f \in \mathscr{C}^\infty(M),\ X \in \mathfrak{X}(M). \tag{102}$$

Then there exists $\alpha \in \Omega^1(M)$ such that $\alpha X = A(X)$ for all $X \in \mathfrak{X}(M)$.

Proof. Let $p \in U$, and let $x_i : U \subset M \to \mathbb{R}$, $i = 1, \ldots, n$, be a coordinate system, with $U \ni p$ an open set. Then any vector field $X \in \mathfrak{X}(M)$ can be written as $X = \sum_i f_i \frac{\partial}{\partial x_i}$ on U, where $f_i \in \mathscr{C}^\infty(U)$. With $h \in \mathscr{C}^\infty(M)$ satisfying $\operatorname{supp} h \subset U$ and $h(p) = 1$, we have

$$X = (1 - h^2)X + h^2X = (1 - h^2)X + \sum_i hf_i \cdot h\frac{\partial}{\partial x_i}. \tag{103}$$

Noting that $hf_i \in \mathscr{C}^\infty(M)$ and $h\frac{\partial}{\partial x_i} \in \mathfrak{X}(M)$, we then infer

$$A(X) = (1 - h^2)A(X) + \sum_i hf_i \cdot A(h\tfrac{\partial}{\partial x_i}). \tag{104}$$

Upon evaluation this at p, we get

$$A(X)_p = \sum_i f_i(p)A(h\tfrac{\partial}{\partial x_i})_p. \tag{105}$$

In particular, if $X_p = 0$ then $f_i(p) = 0$ for all i, and hence $A(X)_p = 0$. In other words, $A(X)_p$ depends only on the value X_p, and the dependence is of course linear. This means that there is a covector $\alpha_p \in T_p^*M$ such that $A(V)_p = \alpha_p V$ for all $V \in T_pM$. As $p \in M$ is arbitrary, the assignment $p \mapsto \alpha_p$ defines a 1-form satisfying

$$\alpha X = A(X) \qquad \text{for all } X \in \mathfrak{X}(M). \tag{106}$$

Smoothness of α is immediate, because $A(X) \in \mathscr{C}^\infty(M)$ for all $X \in \mathfrak{X}(M)$. □

We have seen that given a coordinate system $x_i : U \subset M \to \mathbb{R}$, $i = 1, \ldots, n$, the collections $\{(\mathrm{d}x_i)_p\}$ and $\{(\frac{\partial}{\partial x_i})_p\}$ form a dual system of bases for T_p^*M and T_pM. Thus in U, any vector field X can be written as

$$X = \textstyle\sum_i f_i \dfrac{\partial}{\partial x_i}, \tag{107}$$

with the scalar functions $f_i = (\mathrm{d}x_i)X = X(x_i)$, and any 1-form α can be written as

$$\alpha = \textstyle\sum_i g_i \mathrm{d}x_i \tag{108}$$

with $g_i = \langle \alpha, \frac{\partial}{\partial x_i} \rangle$. At this point, we want to introduce some customary notational conventions for their practicality and built-in error prevention features.

- We will use upper indices (or superscripts) for the coordinate functions: Thus we write $x^i : U \subset M \to \mathbb{R}$ instead of $x_i : U \subset M \to \mathbb{R}$. Then the coordinate-induced dual bases become $\{\mathrm{d}x^i\}$ and $\{\frac{\partial}{\partial x^i}\}$. We also write $\partial_i = \frac{\partial}{\partial x^i}$ if the coordinate system $\{x^i\}$ is clear from the context.
- The components of a vector field will have upper indices, and the components of a 1-form will have lower indices, as in $X = \sum_i f^i \partial_i$ and $\alpha = \sum_i g_i \mathrm{d}x^i$.
- It is convenient to denote the components of a vector field or a 1-form by the same letter as the object itself. Thus we write, e.g., $X = \sum_i X^i \partial_i$ and $\alpha = \sum_i \alpha_i \mathrm{d}x^i$.
- Finally, whenever there is a repeated index in an expression (usually appearing once as a superscript and once as a subscript), summation over that index is assumed, and hence the summation symbol can be omitted. For example, we may write $X = X^i \partial_i$ and $\alpha = \alpha_i \mathrm{d}x^i$. This is called the *Einstein summation convention.* Note that under this convention, if a repeated index appears when summation is not intended, one must make it clear in the context.

With the aforementioned conventions in effect, (107) and (108) become

$$\begin{aligned} X &= X^i \partial_i = \textstyle\sum_i X^i \partial_i = \sum_i \langle \mathrm{d}x^i, X \rangle \partial_i, \quad \text{and} \\ \alpha &= \alpha_i \mathrm{d}x^i = \textstyle\sum_i \alpha_i \mathrm{d}x^i = \sum_i \langle \alpha, \partial_i \rangle \mathrm{d}x^i, \end{aligned} \tag{109}$$

respectively.

Remark 6.3. As a generalization of the coordinate basis $\{\partial_i\}$, we can consider an arbitrary collection $E_i \in \mathfrak{X}(U)$, $i = 1, \ldots, n$, of vector fields, defined on some open set $U \subset M$, such that $\{(E_i)_p\}$ is a basis of T_pM for each $p \in U$. Such a collection is called a *moving frame* on M. A moving frame $\{E_i\}$ induces a unique collection $\theta^i \in \Omega^1(U)$, $i = 1, \ldots, n$, of 1-forms, called the *dual frame* or *coframe*, satisfying

$$\langle \theta^i, E_k \rangle = \delta^i_k, \qquad i, k = 1, \ldots, n, \tag{110}$$

where δ^i_k is the Kronecker delta. Thus $\{(\theta^i)_p\}$ is the basis of T_p^*M dual to $\{(E_i)_p\}$, for each $p \in U$. For any $X \in \mathfrak{X}(U)$ and $\alpha \in \Omega^1(U)$, we have

$$\begin{aligned} X &= X^i E_i = \langle \theta^i, X \rangle E_i, \\ \alpha &= \alpha_i \theta^i = \langle \alpha, E_i \rangle \theta^i. \end{aligned} \tag{111}$$

Now let us consider the frame change

$$F_i = E_k a_i^k, \qquad \omega^i = b_k^i \theta^k, \tag{112}$$

where $A = (a_i^k)$ and $B = (b_k^i)$ are square matrices smoothly depending on $p \in U$. If we think of $\{E_i\}$ as a row vector consisting of vector fields, and $\{\theta^i\}$ as a column vector consisting of 1-forms, then the preceding relation can be written as $F = EA$ and $\omega = B\theta$. The requirement that F and ω are dual to each other yields

$$\delta_k^i = \langle \omega^i, F_k \rangle = b_\ell^i \langle \theta^\ell, E_m \rangle a_k^m = b_\ell^i a_k^\ell, \qquad \text{or} \qquad BA = I. \tag{113}$$

Example 6.4. polar coordinates

Exercise 6.1. Show that the *cotangent bundle* $T^*M = \{(p, \omega) : p \in M, \omega \in T_p^*M\}$ admits a canonical smooth structure, making it a smooth manifold.

7. The exterior derivative

Our next task is to extend the notion of Lie derivatives to operate on 1-forms. The natural operation for 1-forms that replaces push-forward on vector fields is the following: For a smooth map $\phi : M \to N$ and $\alpha \in \Omega^1(N)$, the *pull-back of* α *along* ϕ is a 1-form on M, given by

$$(\phi^*\alpha)_p V = \alpha_{\phi(p)}(\phi_* V), \qquad V \in T_pM. \tag{114}$$

Since $(\phi_*)_p : T_pM \to T_{\phi(p)}N$ is linear for each $p \in M$, it is immediate that the linear map $\phi^*\alpha : \mathfrak{X}(M) \to \mathscr{C}^\infty(M)$ is tensorial and hence $\phi^*\alpha \in \Omega^1(M)$. Note that there is no additional conditions on the map ϕ, in contrast to the case of push-forward for vector fields.

Let $\alpha \in \Omega^1(N)$, let $Y \in \mathfrak{X}(M)$, and let $\phi : M \to N$ be a diffeomorphism. Then $\langle \alpha, \phi_* Y \rangle$ is a scalar function on N, and we have

$$\phi^* \langle \alpha, \phi_* Y \rangle = \langle \alpha, \phi_* Y \rangle \circ \phi = \langle \phi^* \alpha, Y \rangle. \tag{115}$$

Now let Φ_t be the flow of a given vector field $X \in \mathfrak{X}(M)$, and let $\alpha \in \Omega^1(M)$. In light of the preceding identity, we have

$$\Phi_t^* \langle \alpha, Y \rangle - \langle \alpha, Y \rangle = \Phi_t^* \langle \alpha, Y - (\Phi_t)_* Y \rangle + \langle \Phi_t^* \alpha - \alpha, Y \rangle. \tag{116}$$

Dividing both sides by t and sending $t \to 0$ yield

$$\mathfrak{L}_X \langle \alpha, Y \rangle \equiv X \langle \alpha, Y \rangle = \langle \alpha, \mathfrak{L}_X Y \rangle + \langle \mathfrak{L}_X \alpha, Y \rangle, \tag{117}$$

where the last term is defined as follows.

Definition 7.1. In the setting of the preceding paragraph, the *Lie derivative of* $\alpha \in \Omega^1(M)$ *along* $X \in \mathfrak{X}(M)$ is defined by

$$\mathfrak{L}_X \alpha = \frac{d}{dt} (\Phi_t^* \alpha) \Big|_{t=0}. \tag{118}$$

From the product rule (117), we immediately get

$$\langle \mathfrak{L}_X \alpha, Y \rangle = X \langle \alpha, Y \rangle - \langle \alpha, [X, Y] \rangle. \tag{119}$$

This implies that

$$\langle \mathfrak{L}_X \alpha, fY \rangle = X(f \langle \alpha, Y \rangle) - \langle \alpha, X(f)Y + f[X, Y] \rangle = f \langle \mathfrak{L}_X \alpha, Y \rangle, \tag{120}$$

for $f \in \mathscr{C}^\infty(M)$, and hence $\mathfrak{L}_X \alpha \in \Omega^1(M)$. We can also derive the product (or Leibniz) rule

$$\mathfrak{L}_X (f\alpha) = X(f)\alpha + f \mathfrak{L}_X \alpha. \tag{121}$$

Remark 7.2. Recall that for $f \in \mathscr{C}^\infty(M)$ and $X \in \mathfrak{X}(M)$, one has

$$\mathfrak{L}_X f = Xf = \langle \mathrm{d}f, X\rangle. \tag{122}$$

Thus the Lie derivative on functions is the composition of two operations: First the differential produces the 1-form $\mathrm{d}f$, and then this 1-form is applied to the vector field X to produce a scalar function $\langle \mathrm{d}f, X\rangle$. We can write is symbolically as

$$\mathfrak{L}_X = \iota_X \circ \mathrm{d}, \tag{123}$$

where ι_X is the *interior product* (or *contraction*) of a 1-form with the vector field X, defined by $\iota_X\alpha = \langle \alpha, X\rangle$ for $\alpha \in \Omega^1(M)$.

We would like to look for a formula similar to (123) for the Lie derivatives of 1-forms. Let us compute $\mathfrak{L}_X\alpha$ when $\alpha = \mathrm{d}f$ for some scalar function f. Such 1-forms are called *exact 1-forms*. In light of (119), we have

$$\begin{aligned}\langle \mathfrak{L}_X \mathrm{d}f, Y\rangle &= X\langle \mathrm{d}f, Y\rangle - \langle \mathrm{d}f, [X,Y]\rangle = X(Yf) - [X,Y](f)\\ &= Y(Xf) = \langle \mathrm{d}(Xf), Y\rangle = \langle \mathrm{d}(\mathfrak{L}_X f), Y\rangle,\end{aligned} \tag{124}$$

which, upon using (123), implies that

$$\mathfrak{L}_X \circ \mathrm{d} = \mathrm{d} \circ \mathfrak{L}_X = \mathrm{d} \circ \iota_X \circ \mathrm{d}, \tag{125}$$

with both sides understood as operators sending $\mathscr{C}^\infty(M)$ into $\Omega^1(M)$. This means that

$$\mathfrak{L}_X = \mathrm{d} \circ \iota_X, \tag{126}$$

on *exact* 1-forms.

Now we will compute $\mathfrak{L}_X\alpha$ for a general 1-form α. Introducing local coordinates $\{x^i\}$, and writing $\alpha = \alpha_i \mathrm{d}x^i$, we infer

$$\begin{aligned}\mathfrak{L}_X(\alpha_i \mathrm{d}x^i) &= X(\alpha_i)\mathrm{d}x^i + \alpha_i \mathfrak{L}_X(\mathrm{d}x^i) = X(\alpha_i)\mathrm{d}x^i + \alpha_i \mathrm{d}(Xx^i)\\ &= X(\alpha_i)\mathrm{d}x^i + \alpha_i \mathrm{d}X^i,\end{aligned} \tag{127}$$

where $X^i = \langle \mathrm{d}x^i, X\rangle = Xx^i$. On the other hand, we have

$$\mathrm{d}(\alpha X) = \mathrm{d}(\alpha_i \langle \mathrm{d}x^i, X\rangle) = \mathrm{d}(\alpha_i X^i) = X^i \mathrm{d}\alpha_i + \alpha_i \mathrm{d}X^i, \tag{128}$$

and comparing this with the previous equality gives

$$\mathfrak{L}_X\alpha = \mathrm{d}\langle \alpha, X\rangle + X(\alpha_i)\mathrm{d}x^i - X^i \mathrm{d}\alpha_i. \tag{129}$$

If we apply both sides to an arbitrary vector field Y, we get

$$\begin{aligned}\langle \mathfrak{L}_X\alpha, Y\rangle &= \langle \mathrm{d}\langle \alpha, X\rangle, Y\rangle + X(\alpha_i)\langle \mathrm{d}x^i, Y\rangle - X^i\langle \mathrm{d}\alpha_i, Y\rangle\\ &= Y\langle \alpha, X\rangle + X(\alpha_i)Y^i - X^i Y(\alpha_i)\\ &= Y\langle \alpha, X\rangle + \omega(X,Y),\end{aligned} \tag{130}$$

where the last term is defined by

$$\omega(X,Y) = X(\alpha_i)Y^i - Y(\alpha_i)X^i. \tag{131}$$

From a different perspective, by comparing (119) and (130), we infer

$$\omega(X,Y) = X\langle \alpha, Y\rangle - Y\langle \alpha, X\rangle - \langle \alpha, [X,Y]\rangle, \tag{132}$$

which makes it clear that $\omega(X,Y)$ is coordinate independent and defines a bilinear map $\omega : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathscr{C}^\infty(M)$. In fact, one can easily show that ω is tensorial in each of its arguments, and antisymmetric, meaning that

$$\omega(fX,Y) = f\omega(X,Y), \qquad \text{and} \qquad \omega(X,Y) = -\omega(Y,X). \tag{133}$$

Definition 7.3. A bilinear map $\omega : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathscr{C}^\infty(M)$ that is tensorial in each of its arguments is called a *covariant* 2*-tensor on* M. In addition, if ω is antisymmetric, we call it a 2*-form on* M. The space of 2-forms on M is denoted by $\Omega^2(M)$.

If ω is a covariant 2-tensor on M, then in light of (the proof of) Theorem 6.2, for each $p \in M$ there exists a bilinear function $\omega_p : T_pM \times T_pM \to \mathbb{R}$ such that $\omega(X, Y)_p = \omega_p(X_p, Y_p)$. This means that ω has "pointwise values" equal to ω_p, $p \in M$. If ω is a 2-form, then obviously each ω_p must be antisymmetric.

We will consider scalar functions as 0-forms, and set $\Omega^0(M) = \mathscr{C}^\infty(M)$. Thus the differential $\mathrm{d} : \Omega^0(M) \to \Omega^1(M)$ sends 0-forms to 1-forms. The following is its natural extension.

Definition 7.4. The *exterior derivative* $\mathrm{d} : \Omega^1(M) \to \Omega^2(M)$ is defined by

$$(\mathrm{d}\alpha)(X, Y) = X\langle\alpha, Y\rangle - Y\langle\alpha, X\rangle - \langle\alpha, [X, Y]\rangle, \tag{134}$$

for $\alpha \in \Omega^1(M)$ and $X, Y \in \mathfrak{X}(M)$. The differential $\mathrm{d} : \Omega^0(M) \to \Omega^1(M)$ is now renamed to be the exterior derivative on 0-forms.

The interior product $\iota_X : \Omega^1(M) \to \Omega^0(M)$ sends 1-forms to 0-forms. The *interior product* between a 2-form ω and a vector field X is defined by

$$(i_X\omega)(Y) = \omega(X, Y), \qquad Y \in \mathfrak{X}(M). \tag{135}$$

Hence we have $\iota_X : \Omega^2(M) \to \Omega^1(M)$.

Remark 7.5. In terms of the new concepts just defined, (130) can be written as

$$\langle\mathfrak{L}_X\alpha, Y\rangle = Y\langle\alpha, X\rangle + (\mathrm{d}\alpha)(X, Y) = \langle\mathrm{d}\iota_X\alpha, Y\rangle + \langle\iota_X\mathrm{d}\alpha, Y\rangle, \tag{136}$$

yielding

$$\mathfrak{L}_X\alpha = \mathrm{d}\iota_X\alpha + \iota_X\mathrm{d}\alpha, \tag{137}$$

or simply

$$\mathfrak{L}_X = \mathrm{d} \circ \iota_X + \iota_X \circ \mathrm{d}, \tag{138}$$

on 1-forms. This is called *Cartan's formula*, and is the correct extension of (123) to 1-forms.

In the remainder of this section, we shall study the fundamental properties of k-forms (with $k = 0, 1, 2$) and natural operations defined on them. For 0-forms, the exterior derivative is given by $\langle\mathrm{d}f, X\rangle = Xf$. The operator $\mathrm{d} : \Omega^0(M) \to \Omega^1(M)$ is linear, and satisfies

$$\mathrm{d}(fg) = f\mathrm{d}g + g\mathrm{d}f. \tag{139}$$

Furthermore, we have

$$\begin{aligned} (\mathrm{dd}f)(X, Y) &= X\langle\mathrm{d}f, Y\rangle - Y\langle\mathrm{d}f, X\rangle - \langle\mathrm{d}f, [X, Y]\rangle \\ &= X(Yf) - Y(Xf) - [X, Y]f = 0, \end{aligned} \tag{140}$$

meaning that

$$\mathrm{d}^2 = \mathrm{d} \circ \mathrm{d} = 0, \tag{141}$$

on 0-forms. To probe if there is a Leibniz-type rule for the exterior derivative on 1-forms, let $f \in \Omega^0(M)$ and $\alpha \in \Omega^1(M)$, and compute

$$\begin{aligned} (\mathrm{d}(f\alpha))(X, Y) &= X\langle f\alpha, Y\rangle - Y\langle f\alpha, X\rangle - \langle f\alpha, [X, Y]\rangle \\ &= (Xf)\langle\alpha, Y\rangle - (Yf)\langle\alpha, X\rangle + f(\mathrm{d}\alpha)(X, Y) \\ &= \langle\mathrm{d}f, X\rangle\langle\alpha, Y\rangle - \langle\mathrm{d}f, Y\rangle\langle\alpha, X\rangle + f(\mathrm{d}\alpha)(X, Y). \end{aligned} \tag{142}$$

This has the appearance of a Leibniz rule, if we think of the term $\langle\mathrm{d}f, X\rangle\langle\alpha, Y\rangle - \langle\mathrm{d}f, Y\rangle\langle\alpha, X\rangle$ as a 2-form, which is some kind of product between $\mathrm{d}f$ and α, applied to the pair (X, Y). We are thus led to the following product operation between 1-forms.

Definition 7.6. The *exterior product* (or *wedge product*) $\wedge : \Omega^1(M) \times \Omega^1(M) \to \Omega^2(M)$ is the bilinear map defined by

$$(\mu \wedge \nu)(X, Y) = \mu(X)\nu(Y) - \mu(Y)\nu(X), \qquad X, Y \in \mathfrak{X}(M). \tag{143}$$

For $f \in \Omega^0(M)$ and $\alpha \in \Omega^k(M)$, $k = 0, 1, 2$, we set $f \wedge \alpha = \alpha \wedge f = f\alpha$.

It is easy to check that $\mu \wedge \nu$ is indeed a 2-form, and that $\mu \wedge \nu = \nu \wedge \mu$. Moreover, the exterior product is pointwise (or tensorial), in the sense that $(\mu \wedge \nu)_p$ depends only on μ_p and ν_p for $p \in M$. In terms of exterior products, (142) can be rewritten as

$$\mathrm{d}(f \wedge \alpha) = \mathrm{d}f \wedge \alpha + f \wedge \mathrm{d}\alpha. \tag{144}$$

If we switch the order of f and α, we get

$$\mathrm{d}(\alpha \wedge f) = \mathrm{d}(f \wedge \alpha) = \mathrm{d}f \wedge \alpha + f \wedge \mathrm{d}\alpha = -\alpha \wedge \mathrm{d}f + \mathrm{d}\alpha \wedge f. \tag{145}$$

These Leibniz rules, together with (139), can be summarized as

$$\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^k \alpha \wedge \mathrm{d}\beta, \tag{146}$$

for $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^\ell(M)$ with $k + \ell \leq 1$.

Exercise 7.1. Prove the following.

(a) $\mathfrak{L}_X(\alpha \wedge \beta) = \mathfrak{L}_X\alpha \wedge \beta + \alpha \wedge \mathfrak{L}_X\beta$
(b) $\mathfrak{L}_X \circ \mathrm{d} = \mathrm{d} \circ \mathfrak{L}_X$
(c) $\mathfrak{L}_{[X,Y]} = [\mathfrak{L}_X, \mathfrak{L}_Y] := \mathfrak{L}_X \circ \mathfrak{L}_Y - \mathfrak{L}_Y \circ \mathfrak{L}_X$
(d) $\iota_X(\alpha \wedge \beta) = \iota_X\alpha \wedge \beta + (-1)^k \alpha \wedge \iota_X\beta$
(e) $\iota_{[X,Y]} = [\mathfrak{L}_X, \iota_Y] := \mathfrak{L}_X \circ \iota_Y - \iota_Y \circ \mathfrak{L}_X$

Exercise 7.2. Let $\phi : M \to N$ be a smooth map. Prove the following for the pull-back map.

(a) $\phi^*(\alpha \wedge \beta) = \phi^*(\alpha) \wedge \phi^*(\beta)$ for any differential forms α and β on M.
(b) $\phi^*(\mathrm{d}\alpha) = \mathrm{d}(\phi^*\alpha)$ for any differential form α on M.