# MATH 250A: INTRODUCTION TO SMOOTH MANIFOLDS

TODD KEMP

## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>0. Review of Calculus</td>
<td>2</td>
</tr>
<tr>
<td>0.1. Total Derivatives</td>
<td>2</td>
</tr>
<tr>
<td>0.2. Partial and Directional Derivatives</td>
<td>3</td>
</tr>
<tr>
<td>0.3. Taylor's Theorem</td>
<td>4</td>
</tr>
<tr>
<td>0.4. Lipschitz Continuity</td>
<td>6</td>
</tr>
<tr>
<td>0.5. Inverse Function Theorem</td>
<td>9</td>
</tr>
<tr>
<td>0.6. Implicit Function Theorem</td>
<td>10</td>
</tr>
<tr>
<td>0.7. Solutions of ODEs</td>
<td>12</td>
</tr>
<tr>
<td>1. Smooth Manifolds</td>
<td>17</td>
</tr>
<tr>
<td>1.1. Smooth Surfaces in ( \mathbb{R}^d )</td>
<td>17</td>
</tr>
<tr>
<td>1.2. Topological Manifolds</td>
<td>18</td>
</tr>
<tr>
<td>1.3. Smooth Charts and Atlases</td>
<td>20</td>
</tr>
<tr>
<td>1.4. Smooth Structures</td>
<td>21</td>
</tr>
<tr>
<td>2. Smooth Maps</td>
<td>24</td>
</tr>
<tr>
<td>2.1. Definitions, Basic Properties</td>
<td>24</td>
</tr>
<tr>
<td>2.2. Smooth Functions, and Examples</td>
<td>25</td>
</tr>
<tr>
<td>2.3. Diffeomorphisms</td>
<td>27</td>
</tr>
<tr>
<td>2.4. Partitions of Unity</td>
<td>29</td>
</tr>
<tr>
<td>2.5. Applications of Partitions of Unity</td>
<td>32</td>
</tr>
<tr>
<td>3. Tangent Vectors</td>
<td>35</td>
</tr>
<tr>
<td>3.1. Tangent Spaces</td>
<td>35</td>
</tr>
<tr>
<td>3.2. The Differential / Tangent Map</td>
<td>37</td>
</tr>
<tr>
<td>3.3. Local Coordinates</td>
<td>40</td>
</tr>
<tr>
<td>3.4. Velocity Vectors of Curves</td>
<td>43</td>
</tr>
<tr>
<td>3.5. The Tangent Bundle</td>
<td>45</td>
</tr>
<tr>
<td>4. Vector Fields</td>
<td>49</td>
</tr>
<tr>
<td>4.1. Definitions and Examples</td>
<td>49</td>
</tr>
<tr>
<td>4.2. Frames</td>
<td>51</td>
</tr>
<tr>
<td>4.3. Derivations</td>
<td>52</td>
</tr>
<tr>
<td>4.4. Push-Forwards</td>
<td>55</td>
</tr>
<tr>
<td>4.5. Lie Brackets</td>
<td>56</td>
</tr>
<tr>
<td>5. Flows</td>
<td>60</td>
</tr>
<tr>
<td>5.1. Integral Curves</td>
<td>60</td>
</tr>
<tr>
<td>5.2. Flows</td>
<td>62</td>
</tr>
<tr>
<td>5.3. Lie Derivatives</td>
<td>68</td>
</tr>
</tbody>
</table>

Date: December 17, 2014.
0. Review of Calculus

0.1. Total Derivatives.

**Definition 0.1.** Let \( n, m \geq 1 \) be integers, let \( U \subseteq \mathbb{R}^n \) be open, and let \( x_0 \in U \). A function \( f : U \rightarrow \mathbb{R}^m \) is called **differentiable** at \( x_0 \) if there is a linear map \( L : \mathbb{R}^n \rightarrow \mathbb{R}^m \) so that, for sufficiently small \( v \in \mathbb{R}^n \),

\[
f(x_0 + v) - f(x_0) = L(v) + o(|v|).
\]

More precisely, the statement is that

\[
\lim_{v \to 0} \frac{f(x_0 + v) - f(x_0) - L(v)}{|v|} = 0.
\]

When this is true, the linear map \( L \) is called the **derivative** or **total derivative** of \( f \) at \( x_0 \), denoted \( L = Df(x_0) \).

**Proposition 0.2.** Let \( f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m \), and let \( x_0 \in U \). The following facts are easy to verify.

1. If \( f \) is differentiable at a point \( x_0 \), then it is continuous at \( x_0 \).
2. If \( f \) is constant then it is differentiable at all points \( x \), and \( Df(x) = 0 \).
3. If \( f \) is a linear map, then \( f \) is differentiable at all points \( x \), and \( Df(x) = f \); that is, for any base point \( x \), the derivative of \( f \) based at \( x \) is the linear map \([Df(x)]v = f(v)\).

The usual “rules” of differentiation hold for this derivative, as follows.

**Proposition 0.3.** Let \( f, g : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m \) be differentiable at \( x_0 \in U \). Let \( y_0 = f(x_0) \in V \subseteq \mathbb{R}^n \), and let \( h : V \rightarrow \mathbb{R}^p \) be differentiable at \( y_0 \).

1. \( f + g \) is differentiable at \( x_0 \), and \( D(f + g)(x_0) = Df(x_0) + Dg(x_0) \).
2. **(Product Rule)** Suppose \( m = 1 \). Then \( fg \) is differentiable at \( x_0 \), and \( D(fg)(x_0) = f(x_0)Dg(x_0) + g(x_0)Df(x_0) \).
3. **(Chain Rule)** \( h \circ f \) is differentiable at \( x_0 \), and \( D(h \circ f)(x_0) = Dh(y_0) \circ Df(x_0) \).

The proofs of (1) and (2) are left as Exercises; for (3), see [1 Prop C.3, p.643].
0.2. Partial and Directional Derivatives. Given \( f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m \) with \( x_0 \in U \), and a vector \( v \in \mathbb{R}^n \), the directional derivative of \( f \) at \( x_0 \) in the direction \( v \), should it exist, is defined to be

\[
D_v f(x_0) = \lim_{h \to 0} \frac{f(x_0 + hv) - f(x_0)}{h} = \left. \frac{df}{dt} \right|_{t=0}.
\]

As a special case, when we let \( v \) be a standard basis vector, we get the partial derivatives. Note: our standard notation for the coordinates in \( \mathbb{R}^n \) is \((x^1, x^2, \ldots, x^n)\). The standard basis for \( \mathbb{R}^n \) is denoted \( \{e_1, e_2, \ldots, e_n\} \).

**Definition 0.4.** Let \( f : U \subseteq \mathbb{R}^n \to \mathbb{R} \), and let \( x_0 \in U \). For \( 1 \leq j \leq n \), the partial derivative \( \partial f / \partial x^j = \partial_j f \) at \( x_0 \) is the ordinary derivative of \( f \) at \( x_0 \) with respect to \( x^j \), holding all other variables \( x^i \) with \( i \neq j \) constant. That is

\[
\frac{\partial f}{\partial x^j}(x_0) = \lim_{h \to 0} \frac{f(x_0 + he_j) - f(x_0)}{h} = \left. \frac{df}{dt} \right|_{t=0} = D_{e_j} f(x_0),
\]

should this limit exist.

If the partial derivative \( \partial_j f \) exists at each point \( x_0 \in U \), then we can interpret \( \partial_j f \) as a function \( U \to \mathbb{R} \) as well, and ask about its partial derivatives.

**Definition 0.5.** Let \( U \subseteq \mathbb{R}^n \) be open, and let \( k \in \mathbb{N} \). A function \( f : U \to \mathbb{R}^m \) is called \( C^k(U) \) if all mixed partial derivatives of length \( \leq k \):

\[
\frac{\partial^\ell f}{\partial x^{j_1} \cdots \partial x^{j_\ell}}, \quad 1 \leq \ell \leq k, \quad j_1, \ldots, j_\ell \leq n
\]

are continuous functions. If \( f \in C^k(U) \) for all \( k \in \mathbb{N} \), we say \( f \in C^\infty(U) \), and call such a function smooth.

Let \( f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m \) be \( C^1 \). In terms of components \( f = (f^1, \ldots, f^m) \), this means that the \( m \) real-valued functions \( \{f^j : 1 \leq j \leq m\} \) are in \( C^1(U) \). Then we can form the \( m \times n \) matrix

\[
[Jf(x_0)]_j^i = \frac{\partial f^i}{\partial x^j}(x_0).
\]

This is the **Jacobian matrix** of \( f \). (To be clear: \( A_{ij} \) is the entry in the \( i \)th row and \( j \)th column of the matrix \( A \), sometimes denoted \( A_{ij} \).)

**Proposition 0.6.** Let \( f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m \). If \( f \) is differentiable at \( x_0 \in U \), then the partial derivatives \( \partial_j f(x_0), 1 \leq j \leq n \) exists, and in terms of the standard basis, the matrix of \( Df(x_0) \) is \( Jf(x_0) \). Conversely, if \( f \) is \( C^1(U) \), then \( f \) is differentiable at each \( x_0 \in U \), and so again \( Df(x_0) \) has matrix \( Jf(x_0) \).

**Proof.** By definition,

\[
\partial_j f(x_0) = \left. \frac{df}{dt} \right|_{t=0} = D(f \circ \alpha_j)(0),
\]

where \( \alpha_j(t) = x_0 + te_j \) is an affine map \( \mathbb{R} \to \mathbb{R}^n \) with \( \alpha_j(0) = x_0 \). Using Proposition 0.2 we have \( [D\alpha_j(t)](s) = se_j \), and so by the chain rule, this derivative exists and

\[
\partial_j f(x_0) = Df(x_0) \circ D\alpha_j(0).
\]

In particular, we have \( [D\alpha_j(0)](1) = e_j \), and so this tells us that \( \partial_j f(x_0) = [Df(x_0)](e_j) \), which is precisely to say that \( \partial_j f(x_0) \) is the \( j \)th column of the matrix of \( Df(x_0) \) in the standard basis. This shows that \( Jf(x_0) \) is the matrix.
Taylor’s theorem says, in a precise sense, that Notably we have the first and second order expansions

\[ D_v f(x_0) = \frac{d}{dt} f(x_0 + tv) \bigg|_{t=0} = [Df(x_0)](v). \]

In other words: if \( f \) is differentiable, then \( v \mapsto D_v f(x_0) \) is a linear map (it is the total derivative). The linearity of this map is not at all clear from the definition of directional derivative; and, indeed, it is not true for non-differentiable functions (even ones with all partial derivatives existing).

**Corollary 0.7** (Chain rule for partial derivatives). Let \( f = (f_1, \ldots, f^m) \) be a \( C^1 \) function \( U \subseteq \mathbb{R}^n \to \mathbb{R}^m \), let \( x_0 \in U \), and let \( g = (g^1, \ldots, g^p) \) be a \( C^1 \) function \( V \subseteq \mathbb{R}^m \to \mathbb{R}^p \) with \( f(x_0) \in V \). Then for \( 1 \leq i \leq p \)

\[ \partial_j (g^i \circ f)(x_0) = \sum_{k=1}^m \partial_k g^i(f(x_0)) \partial_j f^k(x_0). \]

**Proof.** We simply calculate

\[ [\partial_j (g \circ f)(x_0)]^i = \partial_j (g^i \circ f)(x_0) = [J(g \circ f)(x_0)]^i_j = [D(g \circ f)(x_0)]^i_j = \sum_{k=1}^m [Dg(f(x_0))]^i_k [Df(x_0)]^k_j \]

and this is precisely what is stated above. \( \square \)

**Corollary 0.8.** A composition of smooth functions is smooth.

The proof of Corollary 0.8 is left as an Exercise.

0.3. **Taylor’s Theorem.** We will have frequent use for Taylor approximations of smooth functions. In multivariate form, it is useful to use multi-index notation. Denote \([n] = \{1, \ldots, n\}\). For a positive integer \( k \) and \( J = (j_1, \ldots, j_k) \in [n]^k \), we denote for \( x \in \mathbb{R}^n \)

\[ \partial_J = \partial_j_1 \cdots \partial_j_k, \quad x^J = x^{j_1} x^{j_2} \cdots x^{j_k}. \]

Let \( k \) be a positive integer, and \( f \in C^k(U) \) for some open set \( U \subseteq \mathbb{R}^n \). For \( p \in U \), the \( k \)th order Taylor polynomial of \( f \) at \( p \) is the polynomial function

\[ P_k f(x; p) = \sum_{m=0}^k \frac{1}{m!} \sum_{J \in [n]^m} \partial_J f(p)(x - p)^J. \]

Notably we have the first and second order expansions

\[ P_1 f(x; p) = f(p) + \sum_{j=1}^n \partial_j f(p)(x^j - p^j), \]

\[ P_2 f(x; p) = P_1 f(x; p) + \frac{1}{2} \sum_{i,j=1}^n \partial_i \partial_j f(p)(x^i - p^i)(x^j - p^j). \]

Taylor’s theorem says, in a precise sense, that \( f(x) = P_k f(x; p) + o(|x - p|^k) \). Here we will state and prove this theorem in its most useful version: with integral remainder term.
**Theorem 0.9** (Taylor’s Theorem). Let $U \subseteq \mathbb{R}^n$ be open and convex, let $k$ be a positive integer, and let $f \in C^{k+1}(U)$. Then for any points $x, p \in U$,

$$f(x) = P_k f(x; p) + \frac{1}{k!} \sum_{J \in [n]^{k+1}} (x - p)^J \int_0^1 (1 - t)^k \partial J f(p + t(x - p)) \, dt. \quad (0.1)$$

**Proof.** We prove this by induction on $k$. For $k = 0$, the desired statement is

$$f(x) = f(p) + \sum_{j=1}^n (x^j - p^j) \int_0^1 \partial_j f(p + t(x - p)) \, dt.$$

To see why this is true, let $u(t) = f(p + t(x - p))$. Our assumption here is that $f$ is $C^1$ in a convex neighborhood of $x, p$, and so the function $u$ (defined along the line joining $x$ to $p$) is $C^1[0, 1]$. Hence, by the Fundamental Theorem of Calculus,

$$u(t) - u(0) = \int_0^1 u'(t) \, dt = \int_0^1 [Df(p + t(x - p))](x - p) \, dt$$

and this, together with the fact that $u(0) = f(p)$, yields the result.

Now, suppose (0.1) holds true for a given $k$. Fix a $J \in [n]^{k+1}$. Then, integrating by parts,

$$\int_0^1 (1 - t)^k \partial J f(p + t(x - p)) \, dt$$

$$= \left[ -\frac{(1 - t)^{k+1}}{k + 1} \partial J f(p + t(x - p)) \right]_{t=0}^{t=1} + \int_0^1 \frac{(1 - t)^{k+1}}{k + 1} \frac{\partial}{\partial t} \partial J f(p + t(x - p)) \, dt.$$ 

The first term is simply $\frac{1}{k+1} \partial J f(p)$. For the second term, we compute by the chain rule that that

$$\frac{\partial}{\partial t} (\partial J f)(p + t(x - p)) = \sum_{j=1}^n \partial_j (\partial J f)(p + t(x - p))(x^j - p^j).$$

Thus, summing over $J$, we have

$$\sum_{J \in [n]^{k+1}} (x - p)^J \int_0^1 (1 - t)^k \partial J f(p + t(x - p)) \, dt$$

$$= \frac{1}{k + 1} \sum_{J \in [n]^{k+1}} \partial J f(p)(x - p)^J + \frac{1}{k + 1} \sum_{J \in [n]^{k+1}, j \in [n]} (x - p)^J (x - p)^j \partial_j \partial J f(p + t(x - p))$$

$$= \frac{1}{k + 1} \sum_{J \in [n]^{k+1}} \partial J f(p)(x - p)^J + \frac{1}{k + 1} \sum_{J' \in [n]^{k+2}} (x - p)^J' \partial J' f(p + t(x - p)).$$

But, by the inductive hypothesis, this sum is $k!$ times $f(x) - P_k f(x)$. Dividing out and combining yields (0.1) at stage $k + 1$, concluding the proof. \hfill \Box

**Corollary 0.10.** Let $U$ be an open convex subset of $\mathbb{R}^n$, let $p \in U$, and let $f \in C^{k+1}(U)$ for some positive integer $k$. If all partial derivatives of order $k + 1$ of $f$ are bounded on $U$ — say $|\partial J f(y)| \leq M$ for $J \in [n]^{k+1}$ — then for $x \in U$,

$$|f(x) - P_k f(x)| \leq \frac{n^{k+1}M}{(k + 1)!} |x - p|^{k+1}.$$
Lemma 0.13. Let \( x \) be a radius \( \delta \). Choose a radius \( \delta \).

Proof. By assumption, for each compact, then \( \delta \).

The function is said to be \( \text{Lipschitz} \). Let \( \text{Lipschitz Continuity} \).

Proof. There are \( n^{k+1} \) terms in the sum on the right-hand-side of \( (0.1) \). The \( J \)th term is bounded by

\[
\frac{1}{k!} |(x - p)^J| \int_0^1 (1 - t)^k M dt = \frac{M}{(k + 1)!} |(x^j - p^j) \ldots (x^{j_{k+1}} - p^{j_{k+1}})|.
\]

Now, set \( y = x - p \). The product term is then \( |y^{j_1}| \ldots |y^{j_{k+1}}| \), which can be written uniquely in the form \( |y|^{|j_1|} \ldots |y|^{|j_{n}|} \) for some non-negative integer exponents \( \epsilon_1, \ldots, \epsilon_n \) satisfying \( \epsilon_1 + \cdots + \epsilon_n = k + 1 \). Note that

\[
|y|^{2(k+1)} = (|y|^2 + \cdots + |y|^2)^{k+1} = \sum_{r_1, \ldots, r_n \geq 0} \binom{k + 1}{r_1 \cdots r_n} |y|^{2r_1} \ldots |y|^{2r_n} \geq |y|^{2\epsilon_1} \ldots |y|^{2\epsilon_n}.
\]

Taking square roots shows that \( |y^{j_1}| \ldots |y^{j_{k+1}}| \leq |y|^{k+1} \). Combining this with the above completes the proof. \( \square \)

Taylor’s Theorem is often stated only in the form of Corollary \( \ref{cor:Taylor} \). This is useful for many applications, but it fails to make clear an important regularity result: the terms in the \( o(|x - p|^{k}) \) remainder are all of the form \( (x - p)^J \) (for some \( J \in [n]^{k+1} \)) times a nice function — if the original function is \( C^{k+1+\ell} \), then these term are themselves \( C^{\ell} \). We record this as a proposition.

Proposition 0.11. In Taylor’s Theorem \( \ref{thm:Taylor} \), suppose \( f \in C^{k+1+\ell} \) for some \( \ell > 0 \). Then the remainder terms are \( (x - p)^J \) times functions

\[
x \mapsto \int_0^1 (1 - t)^k \partial_j f(p + t(x - p)) dt
\]

that are \( C^{\ell} \).

Proof. These are functions of the form \( \int_0^1 g(x, t) dt \) where \( g(x, t) = (1 - t)^k \partial_j f(p + t(x - p)) \). Since \( f \in C^{k+1+\ell} \), \( \partial_j f \in C^{\ell} \) for any \( J \in [n]^{k+1} \), and so \( g \) is \( C^{\ell} \) in both \( x \) and \( t \). The result then follows by differentiating with respect to \( x \) (repeatedly) under the integral, which is easily justified in this situation (see, for example, \( \ref{prop:Taylor} \)). \( \square \)

0.4. Lipschitz Continuity.

Definition 0.12. Let \((X, d_X)\) and \((Y, d_Y)\) be metric spaces. A function \( f : X \to X' \) is called \( \text{Lipschitz} \) if there is a constant \( C > 0 \) so that

\[
d_Y(f(x), f(x')) \leq C \cdot d_X(x, x'), \quad \text{for all } x, x' \in X.
\]

The function is said to be \( \text{locally Lipschitz} \) if, given \( x_0 \in X \), there is a neighborhood \( U \) containing \( x_0 \) such that \( f|_U \) is Lipschitz.

Lemma 0.13. Let \( f : X \to Y \) be a locally Lipschitz map between metric spaces. If \( K \subset X \) is compact, then \( f|_K \) is Lipschitz.

Proof. By assumption, for each \( x \in K \) there is neighborhood \( U_x \) of \( x \) on which \( f \) is Lipschitz; choose a radius \( \delta(x) \) so that \( B(x, 2\delta(x)) \subseteq U_x \); then there is a Lipschitz constant \( C(x) \) so that

\[
d_Y(f(x), f(x')) \leq C(x) d_X(x, x'), \quad \text{for all } x, x' \in B(x, 2\delta(x)).
\]
Define there are finitely many points $x_1, \ldots, x_n \in K$ so that $K \subseteq B(x_1, \delta(x_1)) \cup \cdots \cup B(x_n, \delta(x_n))$. Define $C = \max\{C(x_1), \ldots, C(x_n)\}$ and $d = \min\{\delta(x_1), \ldots, \delta(x_n)\}$.

- Suppose $d(x, x') < \delta$. We know there is a points $x_j \in \{x_1, \ldots, x_n\}$ so that $d_X(x, x_j) < \delta(x_j)$. Then by the triangle inequality $d_X(x', x_j) \leq d_X(x, x) + d_X(x, x_j) < \delta + \delta(x_j) \leq 2\delta(x_j)$. Thus, both $x, x'$ are in $U_{x_j}$ where $f$ is Lipschitz with constant $C(x_j) \leq C$, and so $d_Y(f(x), f(x')) \leq C d_X(x, x')$.

- Suppose, on the other hand, that $d_X(x, x') \geq \delta$. Since $f$ is continuous, $f(K)$ is compact (Exercise), and hence it is bounded. Then we have

$$d_Y(f(x), f(x')) \leq \text{diam} f(K) \leq \frac{\text{diam} f(K)}{\delta} d_X(x, x').$$

This shows that $f$ is Lipschitz on $K$, with constant $\leq \max\{C, \text{diam} f(K)/\delta\}$. \qed

The most prevalent examples of Lipschitz functions $\mathbb{R}^n \rightarrow \mathbb{R}^m$ (equipped with their usual Euclidean metrics) are $C^1$ functions.

**Proposition 0.14.** Let $U \subseteq \mathbb{R}^n$ be an open subset, and let $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ be $C^1$. Then $f$ is locally Lipschitz.

**Proof.** Let $x_0 \in U$, and fix any closed ball $\overline{B}$ centered at $x_0$ with $\overline{B} \subseteq U$. Note that $\overline{B}$ is convex, so for any $a, b \in \overline{B}$ the whole line segment between $a$ and $b$ is contained in $\overline{B}$. By the chain rule, $t \mapsto f(a + t(b - a))$ is differentiable, and so by the fundamental theorem of calculus applied to its components, we have

$$f(b) - f(a) = \int_0^1 \frac{d}{dt} f(a + t(b - a)) \, dt.$$

Applying the chain rule, this gives

$$f(b) - f(a) = \int_0^1 [Df(a + t(b - a))] (b - a) \, dt.$$

We can therefore estimate

$$|f(b) - f(a)| \leq \int_0^1 |[Df(a + t(b - a))](b - a)| \, dt.$$

Given any $m \times n$ matrix $A$ and vector $v \in \mathbb{R}^n$, by the Cauchy-Schwarz inequality we have $|Av| \leq |A|_2 |v|$ where $|A|_2$ is the Hilbert-Schmidt / Fröbenius norm of $A$:

$$|A|_2^2 = \sum_{i=1}^m \sum_{j=1}^n |A_{ij}|^2.$$

Thus, we have $|[Df(a + t(b - a))](b - a)| \leq |Df(a + t(b - a))|_2 |b - a|$. By assumption $f \in C^1(U)$, which means that $x \mapsto Df(x)$ is a continuous matrix-valued function, and therefore so is $x \mapsto |Df(x)|_2$. Hence, $M = \max \overline{B} |Df|_2$ exists. Since $|Df(a + t(b - a))|_2 \leq M$, we therefore have

$$|f(b) - f(a)| \leq \int_0^1 |b - a| \, dt = M |b - a|.$$

So $f$ is Lipschitz on the neighborhood $B$ (the interior of $\overline{B}$) of $x_0$, with Lipschitz constant $\leq M$. \qed
The converse of Proposition 0.14 is true in a sense: if \( f \) is locally Lipschitz on an open set in \( \mathbb{R}^n \), then the \( f \) is differentiable almost everywhere.

**Definition 0.15.** A subset \( N \subseteq \mathbb{R}^n \) is said to have measure zero if, for any \( \epsilon > 0 \), there is a countable collection of balls \( \{ B_j \}_{j=1}^\infty \) with \( N \subseteq \bigcup_{j=1}^\infty B_j \), and \( \sum_{j=1}^\infty \text{Vol}(B_j) < \epsilon \).

**Theorem 0.16** (Rademacher). Let \( f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m \) be a locally Lipschitz function. Then \( f \) is differentiable almost everywhere: the set of points in \( U \) at which \( f \) is not differentiable has measure zero.

This theorem is beyond our present purview; the interested reader can find several proofs quickly by Googling.

The best kind of Lipschitz function is one whose Lipschitz constant is \( \leq 1 \).

**Definition 0.17.** A function \( f : X \to Y \) between metric spaces is called a contraction if
\[
d_Y(f(x), f(x')) \leq d_X(x, x'), \quad \text{for all } x, x' \in X.
\]
If there is a constant \( \lambda \in [0, 1) \) such that \( d_Y(f(x), f(x')) \leq \lambda d_X(x, x') \) for all \( x, x' \in X \), then \( f \) is called a strict contraction.

The most basic result for strict contraction is the Banach fixed-point theorem.

**Theorem 0.18** (Banach). Let \( X \) be a non-empty complete metric space. If \( f : X \to X \) is a strict contraction, then \( f \) has a unique fixed-point in \( X \): a unique point \( x \in X \) such that \( f(x) = x \).

**Proof.** First, uniqueness: if \( f(x) = x \) and \( f(y) = y \), then we have
\[
d(x, y) = d(f(x), f(y)) \leq \lambda d(x, y)
\]
and since \( \lambda < 1 \) this is only possible if \( d(x, y) \leq 0 \), meaning \( x = y \).

Now for existence. Fix any initial point \( x_0 \in X \), and consider the sequence of iterates of \( f \):
\[
x_1 = f(x_0), \ x_2 = f(x_1), \ \text{and in general} \ x_n = f(x_{n-1}) = f^n(x_0).
\]
We will show that \( x_n \) converges to a point \( x \), and that this limit is a fixed point of \( f \) (hence the unique fixed point). First, note that
\[
d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq \lambda d(x_n, x_{n-1}).
\]
Iterating this \( n \) times yields
\[
d(x_{n+1}, x_n) \leq \lambda^n d(x_1, x_0).
\]
Thus, by the triangle inequality, for any \( m \geq n \), we have
\[
d(x_n, x_m) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{m-1}, x_m) \leq (\lambda^n + \lambda^{n+1} + \cdots + \lambda^{m-1}) d(x_0, x_1).
\]
The sum is
\[
\sum_{j=n}^{m-1} \lambda^j = \lambda^n \sum_{k=0}^{m-n-1} \lambda^k < \frac{\lambda^n}{1 - \lambda} \to 0 \text{ as } n \to \infty.
\]
This shows that \( (x_n)_{n=0}^\infty \) is a Cauchy sequence. Since \( X \) is a complete metric space, it follows that there is a limit \( x = \lim_{n \to \infty} x_n \). Now, using the continuity of \( f \),
\[
f(x) = f \left( \lim_{n \to \infty} x_n \right) = \lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} x_{n+1} = x.
\]
Thus, \( x \) is a fixed-point of \( f \). \( \square \)
0.5. **Inverse Function Theorem.** Let \( f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m \) be a function differentiable at \( x_0 \). Suppose that \( f \) has a differentiable inverse in a neighborhood of \( x_0 \); so \( f^{-1}(f(x)) = x \) for \( x \) in a neighborhood of \( x_0 \). Denote by \( I_n \) the \( n \times n \) identity matrix. Then by the chain rule

\[
I_n = D(x \mapsto x)(x_0) = D(f^{-1} \circ f)(x_0) = [D(f^{-1})(f(x_0))][Df(x_0)].
\]  

(0.2)

That is: the derivative \( Df(x_0) \) has an inverse matrix (which is \( D(f^{-1})(f(x_0)) \)). A matrix can only be invertible if it is square, and so immediately we see this is only possible if \( n = m \).

A smooth function from an open set \( U \subseteq \mathbb{R}^n \) onto \( V \) which possesses a smooth inverse is called a **diffeomorphism.** The above calculation shows that diffeomorphisms only exists between like-dimensional open sets, and any diffeomorphism has invertible derivative at each point.

The inverse function theorem is a kind of converse to this: if the derivative \( Df(x_0) \) is an invertible matrix, then \( f \) itself has a differentiable inverse in a small neighborhood of \( x_0 \); i.e. \( f \) is a local diffeomorphism. We will state this in the smooth \((C^\infty)\) category, but it holds true for \( C^k \) or just differentiable functions just as well.

**Theorem 0.19** (Inverse Function Theorem). Let \( U \subseteq \mathbb{R}^n \) be open, and let \( f : U \to \mathbb{R}^n \) be smooth. Suppose that, for some point \( x_0 \in U \), \( Df(x_0) \) is invertible. Then there exists a connected neighborhood \( U_0 \subseteq U \) of \( x_0 \) and a connected neighborhood \( V_0 \subseteq f(U_0) \) of \( f(x_0) \), such that \( f|_{U_0} : U_0 \to V_0 \) is a diffeomorphism.

**Proof.** To begin, we rescale and recenter: let \( g(x) = [Df(x_0)]^{-1}(f(x + x_0) - f(x_0)) \); then \( f(x) = [Df(x_0)]g(x - x_0) + f(x_0) \), and so the statement of the theorem is equivalent to showing that \( g \) has an inverse in a connected neighborhood of \( 0 \). Note that \( g(0) = 0 \) and \( Dg(0) = I_n \). Also note: since \( x \mapsto Dg(x) \) is continuous, so is \( x \mapsto \det Dg(x) \); since \( \det Dg(0) = 1 \), there is a neighborhood of \( 0 \) where \( Dg \) is invertible, and so by shrinking the domain of \( g \) (which is \( U + x_0 \) if necessary, we will assume \( Dg \) is invertible everywhere on the domain of \( g \), which we denote \( U' \).

Set \( h(x) = x - g(x) \); then \( Dh(0) = I_n - I_n = 0 \). Since \( x \mapsto Dh(x) \) is continuous, it follows that for all sufficiently small \( x \), \( |Dh(x)|_2 \leq \frac{1}{2} \); so choose \( \delta > 0 \) small enough that \( B(0, \delta) \subseteq U' \) and \( |Dh(x)|_2 \leq \frac{1}{2} \) for \( |x| \leq \delta \). Thus, by Proposition 0.14, if \( x, y \in B(0, \delta) \),

\[
|h(x) - h(x')| \leq \frac{1}{2}|x - x'|.
\]

In particular, taking \( x' = 0 \) and noting that \( h(0) = 0 \), this gives \( |h(x)| \leq \frac{1}{2}|x| \) for \( |x| \leq \delta \).

Now, since \( g(x) - g(x') + h(x) - h(x') = x - x' \), we have

\[
|x - x'| \leq |g(x) - g(x')| + |h(x) - h(x')| \leq |g(x) - g(x')| + \frac{1}{2}|x - x'|, \quad \text{for } x, y \in B(0, \delta).
\]

Rearranging this yields

\[
|x - x'| \leq 2|g(x) - g(x'|. \tag{0.3}
\]

This shows that \( g \) is one-to-one on \( B(0, \delta) \).

**Claim:** given \( y \in B(0, \delta/2) \), there is a unique \( x \in B(0, \delta) \) with \( g(x) = y \). To see this, define \( \tilde{h}(x) = y + h(x) = y + x - g(x) \); thus \( g(x) = y \iff \tilde{h}(x) = x \). So it behooves us to show that \( \tilde{h} \) has a unique fixed point. First note that

\[
|\tilde{h}(x)| \leq |y| + |h(x)| < \frac{\delta}{2} + \frac{1}{2}|x| \leq \delta, \quad \text{for } |x| \leq \delta
\]

and so \( \tilde{h} \) maps the complete metric space \( \overline{B}(0, \delta) \) into itself. We note also that \( |\tilde{h}(x) - \tilde{h}(x')| = |h(x) - h(x')| \leq \frac{1}{2}|x - x'| \), and so \( \tilde{h} \) is a strict contraction. Thus, the claim follows from Theorem 0.18.
Now, set $V_0' = B(0, \delta/2)$ and $U_0' = B(0, \delta) \cap g^{-1}(V_0')$. These are both open sets. The above argument shows that $g : U_0' \to V_0'$ is a bijection. From (0.3) we see that $g^{-1}$ is Lipschitz continuous, and it follows that the preimage $g^{-1}(V_0')$ is also connected. All we have left to do is show that $g^{-1}$ is smooth. If we knew this, (0.2) would show immediately that $D(g^{-1})(g(x)) = [Dg(x)]^{-1}$. In fact, we begin by showing directly that $g^{-1}$ is differentiable, with total derivative given by this formula. That is: fix $y \in V_0'$, and let $x = g^{-1}(y)$. Then for $y' \neq y$ in $V_0'$, setting $x' = g^{-1}(y') \neq x$, we have

$$
\frac{g^{-1}(y') - g^{-1}(y) - [Dg(x)]^{-1}(y' - y)}{|y' - y|} = [Dg(x)]^{-1} \left( \frac{Dg(x)(x' - x) - (y' - y)}{|y' - y|} \right)
$$

Note that $|x' - x||y' - y| \leq 2$ by (0.3). Since $g^{-1}$ is continuous, as $y' \to y$, $x' \to x$. Now, as $x' \to x$, $Dg(x') \to Dg(x)$ since $g$ is $C^1$. Since the matrix function $A \mapsto A^{-1}$ is continuous, it follows that $[Dg(x')]^{-1} \to [Dg(x)]^{-1}$. Since $g$ is differentiable at $x$, the quantity inside the braces tends to 0 as $x' \to x$. All together, this shows that $g^{-1}$ is differentiable at $x$, with derivative $[Dg(x)]^{-1}$.

Thus far, we have shown that $g^{-1}$ is differentiable, and its derivative at $y$ is given by $[Dg(x)]^{-1} = [Dg(g^{-1}(y))]^{-1}$. As a function of $y$, this is a composition:

$$
y \mapsto g^{-1}(y) \xrightarrow{\partial y} Dg(g^{-1}(y)) \xrightarrow{\partial x} [Dg(g^{-1}(y))]^{-1}, \tag{0.4}
$$

where here $Dg$ denotes the map $x \mapsto Dg(x)$, and $\iota(A) = A^{-1}$ is the inverse map on matrices. Each of these functions is continuous (indeed $\iota$ is smooth by, say, Cramer’s rule, which expresses the entries of $A^{-1}$ as rational functions of the entries of $A$), and this shows that the map $y \mapsto D(g^{-1})(y)$ is continuous; hence $g^{-1}$ is $C^1$.

Finally, we show $g^{-1}$ is smooth inductively: suppose we have shown $g^{-1}$ is $C^k$. Then the maps $g^{-1}$ and $Dg$ in (0.4) are $C^k$, and $\iota$ is $C^\infty$, which shows that $D(g^{-1})$ is $C^k$. Hence its components, the partial derivatives of $g^{-1}$, are $C^k$, which is precisely to say that $g^{-1}$ is $C^{k+1}$. Induction now shows that $g \in C^\infty$, concluding the proof. □

### 0.6. Implicit Function Theorem.

An alternative way to view the Inverse Function Theorem is as follows: given $f : \mathbb{R}^n \to \mathbb{R}^n$, consider the function $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ given by $\Phi(x, y) = y - f(x)$. Then the level set $\Phi(x, y) = 0$ can be (globally) described as the graph of $y$ as a function ($f$) of $x$. The Inverse Function Theorem tells us that, if $Df(x_0)$ is invertible, then this level set may also be described (locally) as the graph of $x$ as a function of $y$.

Put in those terms, there is a natural generalization to the case when $n \neq m$.

**Theorem 0.20 (Implicit Function Theorem).** Let $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ be open. Denote the coordinates on $\mathbb{R}^n \times \mathbb{R}^m$ as $(x^1, \ldots, x^n, y^1, \ldots, y^m)$. Let $\Phi : U \to \mathbb{R}^m$ be a smooth function, fix $(x_0, y_0) \in U$, and let $z_0 = \Phi(x_0, y_0)$. Suppose that the $m \times m$ matrix

$$
\begin{bmatrix}
\frac{\partial \Phi_i}{\partial y^j}(x_0, y_0) \\
\end{bmatrix}_{1 \leq i, j \leq m}
$$

is invertible. Then there are neighborhoods $V_0 \subseteq \mathbb{R}^n$ of $x_0$ and $W_0 \subseteq \mathbb{R}^m$ of $y_0$, and a smooth function $f : V_0 \to W_0$ such that the level set $\Phi^{-1}(z_0) \cap (V_0 \times W_0)$ is the graph of $y = f(x)$; that is

for all $(x, y) \in V_0 \times W_0$, $\Phi(x, y) = z_0$ if and only if $y = f(x)$. 

Before giving the proof, we give an example to illustrate the theorem.

**Example 0.21.** Do the equations

\[
\begin{align*}
x^2 - y &= a \\
y^2 - z &= b \\
z^2 - x &= 0
\end{align*}
\]

determine \((x, y, z)\) implicitly as functions of \((a, b)\) in a neighborhood of \((x, y, z, a, b) = (0, 0, 0, 0, 0)\)? To answer, we phrase the question in terms of the level set of a smooth function, and calculate its derivative:

\[
\Phi(x, y, z, a, b) = (x^2 - y - a, y^2 - z - b, z^2 - x), \quad D\Phi(x, y, z, a, b) = \begin{bmatrix}
2x & -1 & 0 & -1 & 0 \\
0 & 2y & -1 & 0 & -1 \\
-1 & 0 & 2z & 0 & 0
\end{bmatrix}.
\]

We want to know if \((x, y, z)\) are locally functions of \((a, b)\); to apply the implicit function theorem, we must consider the 3 \times 3 sub-matrix corresponding to \(\partial_x, \partial_y, \partial_z\), which is

\[
\begin{bmatrix}
2x & -1 & 0 \\
0 & 2y & -1 \\
-1 & 0 & 2z
\end{bmatrix}.
\]

We can readily compute that the determinant of this sub-matrix is \(8xyz - 1\). So, at the point \(x = y = z = a = b = 0\) (which does indeed satisfy the equations), this matrix has determinant \(-1\) which is \(\neq 0\). It therefore follows from the Implicit Function Theorem that, indeed, \((x, y, z)\) can be expressed locally as a function of \((a, b)\) near this point. In fact, this holds true in a neighborhood of any point \((x, y, z, a, b)\) satisfying the equations except possibly when \(xyz = \frac{1}{8}\).

We might ask instead if \((x, z, a)\) can be expressed locally as a function of \((y, b)\). To answer this we need to look at the sub-matrix corresponding to \(\partial_x, \partial_z, \partial_a\), which is

\[
\begin{bmatrix}
2x & 0 & -1 \\
0 & -1 & 0 \\
-1 & 2z & 0
\end{bmatrix}, \quad \text{with determinant constantly equal to 1.}
\]

So this matrix is non-singular at all points, and so indeed \((x, z, a)\) can be expressed locally as a function of \((y, b)\) near any point. Indeed, a little calculation will show that \((x, z, a)\) can be expressed globally as

\[
\begin{align*}
x &= (y^2 - b)^2 \\
z &= y^2 - b \\
a &= (y^2 - b)^4 - y.
\end{align*}
\]

Caution, though: in general, even if the appropriate sub-matrix is always non-singular, the only guarantees that there is a local impact function near each point. It does not generally imply that there is a global implicit function, as there is in this example.

**Proof.** We consider the auxiliary function \(\Psi : U \to \mathbb{R}^n \times \mathbb{R}^m\) defined by \(\Psi(x, y) = (x, \Phi(x, y))\). Then we compute the Jacobian:

\[
D\Psi(x_0, y_0) = \begin{bmatrix}
I_n & \begin{bmatrix}
\frac{\partial \Phi}{\partial x}(x_0, y_0) \\
\frac{\partial \Phi}{\partial y}(x_0, y_0)
\end{bmatrix}_{1 \leq i \leq m}
\end{bmatrix}_{1 \leq i, j \leq m}.
\]
This is a block lower-triangular matrix, with both (square) diagonal blocks invertible (by assumption of the theorem for the lower right block) at \((x, y) = (x_0, y_0)\). It follows that \(D\Psi(x_0, y_0)\) is invertible. It therefore follows from the Inverse Function Theorem that there exist connected neighborhoods \(U_0\) of \((x_0, y_0)\) and \(Y_0\) of \((x_0, z_0)\) such that \(Ψ: U_0 \to Y_0\) is a diffeomorphism. \(U_0\) contains a product neighborhood \(V \times W\) of \((x_0, y_0)\), so shrinking \(U_0\) if necessary and setting \(Y_0 = Ψ(V \times W)\), wlog we assume \(U_0 = V \times W\).

On this neighborhood, we have a local inverse function \(Ψ^{-1}: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m\). We may then write it in the form \(Ψ^{-1}(x, y) = (A(x, y), B(x, y))\), and thus

\[
(x, y) = Ψ(Ψ^{-1}(x, y)) = Ψ(A(x, y), B(x, y)) = (A(x, y), Φ(A(x, y), B(x, y))).
\] (0.5)

Thus \(A(x, y) = x\), and so \(Ψ^{-1}(x, y) = (x, B(x, y))\) for some smooth function \(B: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m\).

Now, let \(V_0 = \{x \in V: (x, z_0) \in Y_0\}\) and \(W_0 = W\), and define \(f: V_0 \to W_0\) as \(f(x) = B(x, z_0)\). Again from (0.5), it follows that

\[
z_0 = Φ(x, B(x, z_0)) = Φ(x, f(x)), \quad \text{for } x \in V_0.
\]

This shows that the graph of \(f\) is indeed contained in \(Φ^{-1}(z_0)\). Conversely, let \((x, y) \in V_0 \times W_0\) be any point where \(Φ(x, y) = z_0\). Then \(Ψ(x, y) = (x, Φ(x, y)) = (x, z_0)\), and so

\[
(x, y) = Ψ^{-1}(x, z_0) = (x, B(x, z_0)) = (x, f(x)),
\]

showing that \(y = f(x)\). So any such point is in the graph of \(f\). This completes the proof. \(\square\)

0.7. Solutions of ODEs. An initial value problem is a system of ODEs (ordinary differential equations) of the following form: given an open interval \(J \subseteq \mathbb{R}\), an open subset \(U \subseteq \mathbb{R}^n\), a continuous function \(F: J \times U \to \mathbb{R}^n\), an initial time \(t_0 \in J\), and an initial condition \(y_0 \in U\), we seek an \(\mathbb{R}^n\)-valued differentiable function \(y\) defined on (a subset of) \(J\) satisfying

\[
\dot{y}(t) = F(y_0, y(t)), \quad y(t_0) = y_0.
\]

Note: \(\dot{y}(t) = y'(t) = \frac{dy}{dt}\); the notation \(\dot{y}\) is convenient when there are superscript around, and is common (especially in physics) to denote a time derivative, which will be our perspective in this course.

The fundamental theorem, stated below, is that if the driving vector field \(F\) is sufficiently smooth (at least locally Lipschitz), then there always exists a unique solution to the initial value problem for any given initial data \(x_0\). (The solution may only exist for a short time.) Moreover, the dependence of the solution on the initial data is as smooth as the driving field. This last point is not often covered in ODE courses (because it is much harder to prove), but it will be very important to us in this course.

The special case in which the driving field \(F\) does not depend explicitly on time \(t\) is called the autonomous case. It turns out the general case can be reduced to the autonomous case, as we will see later; for now, we restrict to autonomous systems.

**Theorem 0.22** (Picard–Lindelöf Theorem for Autonomous ODEs). Suppose \(U \subseteq \mathbb{R}^n\) and \(J \subseteq \mathbb{R}\) are open, and \(F: U \to \mathbb{R}^n\) is locally Lipschitz. Fix a point \((t_0, y_0) \in J \times U\), and consider the initial value problem

\[
\dot{y}(t) = F(y(t)), \quad y(t_0) = y_0.
\] (ODE)

1. **Local Existence**: There exists an open interval \(J_0\) containing \(t_0\), and a \(C^1\) function \(y: J_0 \to U\) that solves (ODE).
(2) **UNIQUENESS:** Any two differentiable solutions to (ODE) agree on the intersection of their domains.

(3) **SMOOTH DEPENDENCE ON INITIAL CONDITIONS:** Let \( J_0 \subseteq \mathbb{R} \) be an open interval containing \( t_0 \) and let \( U_0 \subseteq U \) be open. Suppose that \( \theta : J_0 \times U_0 \to U \) is a function such that \( y(t) = \theta(t, y_0) \) is a solution to (ODE) with initial condition \( y_0 \). Then \( \theta \) is as smooth as \( F \) (i.e. if \( F \in C^k(U) \) then \( \theta \in C^k(J_0 \times U_0) \)).

**Remark 0.23.** As we will see in the proof, a priori the interval \( J_0 \) on which the solution \( y \) exists depends on the initial condition \( x_0 \); however, this interval depends on \( x_0 \) in a smooth way if the boundary of \( U \) is smooth. In this common kind of situation, it then makes sense for the function \( \theta \) in (3) to be defined on a uniform interval \( J_0 \) for varying initial conditions, as we explicitly assume there.

An important technical tool in the proof of Theorem 0.22 is the following ODE comparison theorem (which is a generalized version of Gronwall’s inequality).

**Lemma 0.24** (Gronwall’s Lemma). Let \( J \subseteq \mathbb{R} \) be an open interval, and let \( f : [0, \infty) \to [0, \infty) \) be Lipschitz. Suppose \( u : J \to \mathbb{R}^n \) is a differentiable function which satisfies the differential inequality

\[
|\dot{u}(t)| \leq f(|u(t)|), \quad \text{for all } t \in J.
\]

Fix \( t_0 \in J \), and suppose \( v : [0, \infty) \to [0, \infty) \) is continuous, and differentiable on \( (0, \infty) \), and satisfies the initial value problem

\[
\dot{v}(t) = f(v(t)), \quad v(0) = |u(t_0)|.
\]

Then, for all \( t \in J \), it follows that \( |u(t)| \leq v(|t - t_0|) \).

**Proof.** First, by shifting \( \tilde{u}(t) = u(t + t_0) \) and \( \tilde{J} = J - t_0 \), we may wlog take \( t_0 = 0 \) (and so we rename \( \tilde{J} \) to \( J \) and \( \tilde{u} \) to \( u \)). Let \( J^+ = J \cap [0, \infty) \). If \( t \) is such that \( u(t) = 0 \), then the desired inequality \( |u(t)| \leq v(|t|) \) holds trivially as \( v \geq 0 \); so we restrict our attention to the set \( \{ t \in J^+ : u(t) > 0 \} \). Since \( u \) is differentiable, hence continuous, this is an open set; on this set, \( |u(t)| \) is also differentiable. We can then calculate

\[
\frac{d}{dt} |u(t)| = \frac{1}{|u(t)|} u(t) \cdot \dot{u}(t) \leq \frac{1}{|u(t)|} |u(t)||\dot{u}(t)| = |\dot{u}(t)| \leq f(|u(t)|),
\]

(0.6)

where the first inequality is the Cauchy-Schwarz inequality, and the second is an assumption of the lemma.

Now, let \( A \) be a Lipschitz constant for \( f \), and define \( w(t) = e^{-At}(|u(t)| - v(t)) \). Then \( w(0) = 0 \), and for \( t \in J^+ \), the desired conclusion \( |u(t)| \leq v(t) \) is equivalent to \( w(t) \leq 0 \). Well, if \( t \in J^+ \) and \( w(t) > 0 \) (meaning that \( |u(t)| > v(t) \geq 0 \)), it follows that \( w \) is differentiable, and we can calculate

\[
\dot{w}(t) = -Ae^{-At}(|u(t)| - v(t)) + e^{-At} \frac{d}{dt} (|u(t)| - v(t))
\]

\[
\leq -Ae^{-At}(|u(t)| - v(t)) + e^{-At} (f(|u(t)|) - f(v(t)))
\]

which follows from (0.6) along with the assumption \( \dot{v}(t) = f(v(t)) \). But \( A \) is a Lipschitz constant for \( f \), and so

\[
f(|u(t)|) - f(v(t)) \leq |f(|u(t)|) - f(v(t))| \leq A|u(t)| - v(t) = A(|u(t)| - v(t)).
\]

This shows that \( \dot{w}(t) \leq 0 \).
Since \( w(0) = 0 \), if \( w \) were differentiable for all \( t \in J^+ \) we could now conclude that \( w(t) \leq 0 \) which would conclude the proof. As it stands, we need to be a little more careful. Suppose, for a contradiction, that there is some time \( t_1 \in J^+ \) where \( w(t_1) > 0 \). Define
\[
\tau = \sup\{ t \in [0, t_1] : w(t) \leq 0 \}.
\]
Then, since \( w \) is continuous, \( w(\tau) = 0 \), and by definition \( w(t) > 0 \) for \( \tau < t < t_1 \). But then \( w \) is differentiable on \((\tau, t_1)\), and continuous on \([\tau, t_1]\), and so by the Mean Value Theorem there is some \( t \in (\tau, t_1) \) where \( \dot{w}(t) > 0 \), even though \( w(t) > 0 \), contradicting the above calculations showing that if \( w(t) > 0 \) for \( t \in J^+ \) then \( \dot{w}(t) \leq 0 \). Thus, we have shown that \( w(t) \leq 0 \) for \( t \in J^+ \).

To conclude, we simply replace \( t \) by \(-t\) in the above argument to show the result also holds for \( t \in J^- = J \cap (-\infty, 0] \).

Let us now proceed to the proof of Theorem \ref{thm:0.22}. We begin with the easy case: uniqueness.

**Proof of Theorem \ref{thm:0.22} (2).** Let \( y_1, y_2 \) be two solutions of \eqref{ode} defined on the same interval \( J_0 \), but with potentially different initial conditions. Shrink the interval slightly to an open interval \( J_1 \) containing \( t_0 \) such that \( J_1 \subset J_0 \). The union of the two continuous paths \( y_2(J_1) \) and \( y_2(J_1) \) is a compact subset of \( U \). So, by Lemma \ref{lem:0.13}, the locally Lipschitz vector field \( F \) is globally Lipschitz on this union of paths; let \( C \) be a Lipschitz constant. Thus
\[
\left| \frac{d}{dt}(y_1(t) - y_2(t)) \right| = |F(y_1(t)) - F(y_2(t))| \leq C|y_1(t) - y_2(t)|.
\]
Applying Lemma \ref{lem:0.24} with \( u(t) = y_1(t) - y_2(t), f(v) = Cv, \) and \( v(t) = e^{Ct}|y_1(t_0) - y_2(t_0)| \) yields
\[
|y_1(t) - y_2(t)| \leq e^{C|t-t_0|}|y_1(t_0) - y_2(t_0)|, \quad t \in J_1.
\]
Thus, if the initial conditions \( y_1(t_0) = y_2(t_0) \) are the same, then \( y_1(t) = y_2(t) \) for all \( t \in J_1 \). Since every point of \( J_0 \) is contained in some closed subinterval \( J_1 \) of \( J \), the result now follows. \( \square \)

No we continue with the existence proof.

**Proof of Theorem \ref{thm:0.22} (1).** By assumption \( F \) is locally Lipschitz on \( U \). We begin by restricting to some closed ball \( \overline{B} = \overline{B}_r(y_0) \subset U \) that contain \( y_0 \) (we choose \( r \) small enough that this closed ball is contained in \( U \)); by Lemma \ref{lem:0.13}, \( F \) is Lipschitz on \( \overline{B} \).

They key idea is to rewrite \eqref{ode} as an integral equation:
\[
y(t) = x_0 + \int_{t_0}^t F(y(s)) \, ds.
\]
Eq. \eqref{ode} is equivalent to \eqref{ode} by the fundamental theorem of calculus. Indeed, if \( y \) is any solution to \eqref{ode}, since \( F \) is (assumed) continuous, it follows from \eqref{ode} that \( y \in C^1(J_0) \), and so the fundamental theorem of calculus applies. Conversely, if \( y : J_0 \to U \) is any continuous function satisfying \eqref{ode}, again the continuity of \( F \) shows that, by the fundamental theorem of calculus, \( y \) satisfies \eqref{ode} (and therefore is \( C^1 \) as above). Henceforth, we show that \eqref{ode} has a solution.

We now introduce a linear operator on the space \( C^0(J_0; \overline{B}) \) of continuous maps \( J_0 \to U \):
\[
(Iy)(t) = x_0 + \int_{t_0}^t F(y(s)) \, ds.
\]
(Note that we have not yet defined the interval $J_0$; for the time being, the definition depends on an arbitrarily chosen open interval $J_0 \subseteq J$.) For any continuous $y$, as $F$ is continuous, the integrand of $Iy$ is continuous, and hence (by the fundamental theorem of calculus) $Iy \in C^1(J_0; \mathbb{R}^n) \subset C^0(J_0; \mathbb{R}^n)$. That is,

$$I : C^0(J_0; \overline{B}) \to C^0(J_0; \mathbb{R}^n)$$

is a linear operator. Note that any solution of (0.7) is a fixed point for $I$, and vice versa. To prove such a fixed point exists (and conclude the proof), we want to put a metric space structure on $C^0(J_0; U)$ and show that $I$ is a strict contraction, to use the Banach fixed point theorem. The delicate point is that $I$ does not a priori map $C^0(J_0; U)$ back into itself. This is where the choice of $J_0$ comes in: we can make it small enough that $I$ does map it into itself, simply because $F$ is bounded. Let $M = \sup_{\overline{B}} |F|$, and choose $0 < \epsilon < r/M$. We will take $J_0 = (t_0 - \epsilon, t_0 + \epsilon)$. Then we have the following.

**Claim.** The linear operator $I$ maps $C^0(J_0; \overline{B})$ into itself. To see this, since we’ve already discussed above that $I$ maps continuous functions to continuous functions, it suffices to show that $Iy(t) \in \overline{B}$ for any $t \in J_0$. To that end, we estimate as follows:

$$|Iy(t) - y_0| = \left| y_0 + \int_{t_0}^t F(y(s)) \, dx - y_0 \right| \leq \int_{t_0}^t |F(y(s))| \, ds \leq \int_{t_0}^t M \, ds < M \epsilon \leq r.$$

This proves the claim.

Now we must introduce a metric on $C^0(J_0; \overline{B})$ and show that $I$ is a strict contraction, which will allow us to use the Banach Fixed Point Theorem to prove that $I$ has a unique fixed point, concluding the proof. There is a standard metric for continuous functions on a compact set, given by the sup-norm: $\|y\|_{\infty} \equiv \sup_{J_0} |y|$. Thus, we define

$$d_{\infty}(y_1, y_2) \equiv \sup_{t \in J_0} |y_1(t) - y_2(t)|.$$

It is a standard result that this is a complete metric on $C^0(J_0; \overline{B})$ if $J_0$ is compact; in our case, we have functions $y$ that are continuous on a bigger interval $J$. We are taking $J_0 = (t_0 - \epsilon, t_0 + \epsilon)$ where (so far) we know $\epsilon < r/M$. We must also choose $\epsilon$ small enough that the closed interval $[t_0 - \epsilon, t_0 + \epsilon]$ is contained in $J$. Then we know that our solution is in $C^0(\overline{J_0}; \overline{B})$, which is a complete metric space in $d_{\infty}$. (Note: completeness is just the statement that any uniformly Cauchy sequence of continuous functions has a continuous limit.)

So, we are left to show that $I$ is a strict contraction on the metric space $C^0(J_0; \overline{B})$. To guarantee this, we need one more constraint on $\epsilon$: Let $C$ be a Lipschitz constant for $F$ on $\overline{B}$; then we make sure that $\epsilon < 1/C$. Then we have for any two $y_1, y_2 \in C^0(J_0; \overline{B})$,

$$d_{\infty}(Iy_1, Iy_2) = \sup_{t \in J_0} \left| \int_{t_0}^t F(y_1(s)) \, ds - \int_{t_0}^t F(y_2(s)) \, ds \right| \leq \sup_{t \in J_0} \int_{t_0}^t |F(y_1(s)) - F(y_2(s))| \, ds \leq \sup_{t \in J_0} \int_{t_0}^t C|y_1(s) - y_2(s)| \, ds \leq C \cdot \sup_{t \in J_0} \int_{t_0}^t d_{\infty}(y_1, y_2) \, ds = C|t - t_0|d_{\infty}(y_1, y_2) < C \epsilon \cdot d_{\infty}(y_1, y_2).$$

Since we choose $\epsilon < 1/C$, the constant $C \epsilon$ is strictly less than 1, and hence, by the Banach fixed point theorem, $I$ possesses a unique fixed point in $C^0(J_0; \overline{B})$. As discussed above, this fixed point is the desired solution to (ODE).
To summarize: we have shown that for any \( \epsilon > 0 \) small enough that \( J_0 = (t_0 - \epsilon, t_0 + \epsilon) \) is contained in \( J \), and satisfying \( C_\epsilon < 1 \) and \( M_\epsilon < r \) (where \( C \) is a Lipschitz constant for \( F \), \( M \) is the maximum of \( F \), and \( r \) is the distance from \( y_0 \) to the boundary of \( U \)), it follows that \((\text{ODE})\) has a solution which exists on the interval \( J_0 \).

**Remark 0.25.** This proof is often called “Picard iteration”: we showed there is a solution by appealing to the Banach fixed point Theorem [0.18]. The proof of that theorem is by iteration. So, in total, the above proof shows that the solution to \((\text{ODE})\) can be achieved by iteration of the operator \( I \) starting at any continuous function: i.e. the solution is given by \( y(t) = \lim_{n \to \infty} I^n(y_0)(t) \) (where, for convenience, we choose our starting point to be the constant function \( y = y_0 \)).

For extra nerdiness, we will refer to this proof as “the Picard maneuver”. :)

As for Theorem [0.22][3], which is an important result for us: the proof is quite involved and technical, so we will leave it for the time being. The interested reader may find it on [1, pp. 667-672].

We conclude this section by showing how the general (non-autonomous) case follows as a special case of the autonomous case (in one dimension higher).

**Theorem 0.26** (Full Picard–Lindelöf Theorem). Suppose \( U \subseteq \mathbb{R}^n \) and \( J \subseteq \mathbb{R} \) are open, and \( F: J \times U \to \mathbb{R}^n \) is locally Lipschitz. Fix a point \((t_0, y_0) \in J \times U\), and consider the initial value problem

\[
\dot{y}(t) = F(t, y(t)), \quad y(t_0) = y_0. \tag{ODE'}
\]

1. **Local Existence:** There exists an open interval \( J_0 \) containing \( t_0 \), and a \( C^1 \) function \( y: J_0 \to U \) such that \((\text{ODE'})\) holds.
2. **Uniqueness:** Any two differentiable solutions to \((\text{ODE})\) agree on the intersection of their domains.
3. **Smooth Dependence on Initial Conditions:** Let \( J_0 \subseteq \mathbb{R} \) be an open interval containing \( t_0 \) and let \( U_0 \subseteq U \) be open. Suppose that \( \theta: J_0 \times U_0 \to U \) is a function such that \( y(t) = \theta(t, y_0) \) is a solution to \((\text{ODE})\) with initial condition \( y_0 \). Then \( \theta \) is as smooth as \( F \) (i.e. if \( F \in C^k(U) \) then \( \theta \in C^k(J_0 \times U_0) \)).

**Proof.** Let \( F = (1, F) \) be the \( \mathbb{R}^{n+1} \)-valued vector field whose first component (which we label \( F^0 \) for convenience) is constantly 1; this vector field is locally Lipschitz since \( F \) is. Consider the following autonomous ODE in \( \mathbb{R}^{n+1} \):

\[
\dot{y}(t) = F(y(t)) = (1, F(y^0(t), y(t))), \quad y(t_0) = (t_0, y_0). \tag{0.8}
\]

The first component of this equation is simply \( y^0(t) = t \) with initial condition \( y(t_0) = t_0 \), whose unique solution is \( y(t) = t \). Hence, given the unique solution \( y = (y^0, y) \) of \((0.8)\) (guaranteed to exist by Theorem [0.22][1]), the last \( n \) components \( y \) give a solution to \((\text{ODE'})\), which is unique by Theorem [0.22][2] (since uniqueness of \( y \) clearly implies uniqueness of \( y \)), and depends as smoothly as \( F \) on initial conditions by Theorem [0.22][3] since \( F \in C^k \) iff \( F \in C^k \).
1. Smooth Manifolds

1.1. Smooth Surfaces in \( \mathbb{R}^d \). A smooth manifold is a generalization of a classical smooth surface. The most basic kind of surface is the graph of a smooth function. That is: if \( f: U \subseteq \mathbb{R}^n \to \mathbb{R}^m \) is smooth, then the graph \( \Gamma(f, U) = \{(x, y) \in U \times \mathbb{R}^m : y = f(x)\} \) is a smooth surface. What does this mean? Consider the associated function \( \psi: U \to \mathbb{R}^n \times \mathbb{R}^m \) given by

\[
\psi(x) = (x, f(x)).
\]

By definition, \( \psi \) is a bijection between \( U \) and \( \Gamma(f, U) \). In fact, it is more than a bijection: it is a homeomorphism. The set \( \Gamma(f, U) \) is a (topological) subspace of \( \mathbb{R}^n \times \mathbb{R}^m \), and it inherits a topology from this imbedding; the function \( \psi \) is a continuous bijection from \( U \) onto this space \( \Gamma(f, U) \), and its inverse is also continuous. (This is left as an exercise.) So, the surface \( \Gamma(f, U) \) is, topologically, the same as the open set \( U \) in \( \mathbb{R}^n \).

Not every surface is the graph of a function, globally.

Example 1.1. The \( n \)-sphere is the subset of \( \mathbb{R}^{n+1} \) of points unit distance from the origin: \( S^n = \{x \in \mathbb{R}^{n+1} : |x| = 1\} \). It is not the graph of any function of any variable \( x^j \) in terms of the other \( n \) variables, as any coordinate line \( x^j = c \) with \( |c| < 1 \) passes through \( S^n \) in two points. In terms of coordinates, \( S^n \) is the set of points \((x^1, \ldots, x^{n+1})\) such that \((x^1)^2 + \cdots + (x^{n+1})^2 = 1\); that is, \( S^n \) is the level set of the function \( \Phi(x^1, \ldots, x^{n+1}) = (x^1)^2 + \cdots + (x^{n+1})^2 \) at height 1. Now, \( D\Phi(x) = 2[\{x^1, \ldots, x^{n+1}\}] \). This row matrix is never 0 for \( x \in \mathbb{R}^n \), which means that it is always full rank. Thus, at every point \( x_0 \in S^n \), we can find some coordinate \( x^j \) such that \( \frac{\partial \Phi}{\partial x_j} \neq 0 \), and by the Implicit Function Theorem, there exists a smooth local function \( f: V \subseteq \mathbb{R}^n \to \mathbb{R} \) so that, near \( x_0 \), \( S^n \) is the graph of \( x^j = f(x^1, \ldots, x^{j-1}, x^{j+1}, \ldots, x^n) \). Thus, even though \( S^n \) is not the graph of a smooth function globally, it is locally.

This motivates our more general definition.

Definition 1.2. Let \( n < d \) be positive integers. A smooth \( n \)-dimensional surface in \( \mathbb{R}^d \) is a set \( S \subseteq \mathbb{R}^d \) with the property that, for each point \( x_0 \in S \), there is an open set \( V_{x_0} \subseteq \mathbb{R}^n \), and a smooth function \( f: V_{x_0} \to \mathbb{R}^{d-n} \), and some partition of coordinates \( v = (x^{j_1}, \ldots, x^{j_n}) \) and \( w = (x^{k_1}, \ldots, x^{k_{d-n}}) \) (where \( \{j_1, \ldots, j_n, k_1, \ldots, k_{d-n}\} = \{1, \ldots, d\} \)), such that \( x \in S \) with \( x \sim (v, w) \) and \( v \in V_{x_0} \) if and only if \( w = f(v) \). In other words, a smooth \( n \)-dimensional surface in \( \mathbb{R}^d \) is a subset which is locally the graph of \( d-n \) variables as functions of the other \( n \).

This definition is set up precisely to conform to the Implicit Function Theorem. Indeed, let us make one more definition.

Definition 1.3. Let \( \Phi: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m \) be a smooth function. A point \( c \in \mathbb{R}^m \) is called a regular value for \( \Phi \) if, for any \( x \) in the level set \( \Phi^{-1}(c) \), the derivative \( D\Phi(x) \) has full rank \( m \).

Corollary 1.4. If \( \Phi: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m \) is a smooth function and \( c \in \mathbb{R}^m \) is a regular value of \( \Phi \), then the level set \( \Phi^{-1}(c) \) is a smooth \( n \)-dimensional surface in \( \mathbb{R}^{n+m} \).

Proof. Choose \( m \) pivotal columns in the matrix \( D\Phi(x) \); then the \( m \times m \) submatrix composed of these columns is invertible. The Implicit Function Theorem therefore implies that the level set \( \Phi^{-1}(c) \) is locally the graph of a function of these \( m \) variables in terms of the remaining \( n \) variables. This is the definition of a smooth \( n \)-dimensional surface.

Example 1.5. The Orthogonal Group \( O(n) \) consists of those \( n \times n \) real matrices \( Q \) with the property that \( Q^\top Q = I_n \) (i.e. invertible matrices with \( Q^{-1} = Q^\top \)). Note that the condition is that, if \( Q \) is
the $i$th column of $Q$, then $[Q^T Q]_j^i = \langle Q_i, Q_j \rangle = \delta_j^i$; so the columns of $Q$ form an orthonormal basis for $\mathbb{R}^n$. Of course, since $Q^{-1} = Q^T$ we also have $QQ^T = I_n$, and so the rows of $Q$ also form an orthonormal basis for $\mathbb{R}^n$. It is a group under matrix multiplication: if $Q_1, Q_2 \in O(n)$, then $(Q_1 Q_2)^{-1} = Q_2^{-1} Q_1^{-1} = Q_1^T Q_2^T$; and $(Q^{-1})^T = (Q^T)^{-1} = Q$, so $O(n)$ is closed under product and inverse (and clearly contains the identity $I_n$). Note that, if $x, y \in \mathbb{R}^n$ are vectors, and $Q \in O(n)$, then $\langle Qx, Qy \rangle = \langle Q^T Q x, y \rangle = \langle x, y \rangle$; that is, $O(n)$ consists of linear isometries of $\mathbb{R}^n$. In fact, it is the group of all linear isometries of $\mathbb{R}^n$.

By definition, $O(n)$ is the level set $\Phi^{-1}(I_n)$ where $\Phi(A) = A^\top A$. The function $\Phi$ is a smooth map from $M_n \rightarrow M_n$ (indeed, all its entries are polynomial functions); here $M_n$ is the $n^2$-dimensional vector space of $n \times n$ real matrices. It is better to view this map having a smaller codomain: the set $M_n^{sa}$ of self-adjoint (i.e. symmetric) matrices, since $\Phi(A)^\top = (A^\top A)^\top = A^\top A = \Phi(A) \in M_n^{sa}$. It is easy to check that the dimension of $M_n^{sa}$ is $\frac{n(n+1)}{2}$. So we have a smooth function

$$\Phi: M_n \rightarrow M_n^{sa}$$

and our set of interest is the level set $O(n) = \Phi^{-1}(I_n)$. Now, let us compute the derivative of $\Phi$. For any given $A \in M_n$, $D\Phi(A): M_n \rightarrow M_n^{sa}$ is the linear map given by

$$[D\Phi(A)](H) = A^\top H + H^\top A.$$ 

(Showing this is the case, from the definition of the derivative, is left as an exercise.) Now, consider this derivative at any point $Q \in O(n)$. In fact, if $X$ is any symmetric matrix $X \in M_n^{sa}$, then by taking $H = \frac{1}{2} Q X$, note that

$$[DF(Q)]\left(\frac{1}{2} X Q \right) = \frac{1}{2} Q^\top Q X + \frac{1}{2} X^\top Q Q = \frac{1}{2} (X + X^\top) = X.$$ 

This shows that the linear map $D\Phi(Q): M_n \rightarrow M_n^{sa}$ is surjective, i.e. full-rank, for any $Q \in O(n)$. Thus, $I_n$ is a regular value of $\Phi$, and so by Corollary [1.4] $O(n)$ is a smooth surface in $M_n$, whose dimension is $n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$.

1.2. Topological Manifolds. A smooth manifold is a generalization of a smooth surface. In fact, we will eventually see that it is no generalization at all: every $n$-dimensional smooth manifold is a smooth surface in $\mathbb{R}^d$ for some $d \leq 2n$. However, we do not want to be biased by the imbedding of our manifold, and instead view it as an abstract space. Before we get to smooth manifolds, we must first define (topological) manifolds.

**Definition 1.6.** An $n$-dimensional manifold $M = M^n$ is a second-countable Hausdorff space that is locally Euclidean. To be precise: $M$ is

- **Second-Countable**: there is a countable collection $\mathcal{U} = \{ U_n \}_{n \in \mathbb{N}}$ of open sets in $M$ with the property that any open set in $M$ is a union of elements of $\mathcal{U}$.
- **Hausdorff**: given any two distinct points $p, q \in M$, there are two open sets $U, V \subset M$ so that $p \in U, q \in V$, and $U \cap V = \emptyset$.
- **Locally Euclidean**: given any $p \in M$, there is an open neighborhood $U \subseteq M$ of $p$ that is homeomorphic to an open set in $\mathbb{R}^n$. Let $\hat{U}$ be such an open set in $\mathbb{R}^n$, and let $\varphi: U \rightarrow \hat{U}$ be a homeomorphism. The pair $(U, \varphi)$ is called a coordinate chart near $p$.

The second-countability condition is a requirement to prevent exotic topological spaces that are "too big" from being manifolds. The Hausdorff condition is there to prevent us from including exotic, pathological spaces (like the real line with two origins) from being manifolds. These are technical conditions; the heart of the definition is (3), that $M$ is locally Euclidean.
Proposition 1.7 (Smooth surfaces are manifolds). If $S$ is an $n$-dimensional smooth surface in $\mathbb{R}^d$, then $S$ in an $n$-dimensional manifold.

Proof. Since $\mathbb{R}^d$ is second-countable and Hausdorff, so is any subspace, including $S$. Now, from the definition of smooth surface, we know that, for any $x_0 \in S$, there is a partition of coordinates $x \sim (v, w)$ with $v \in \mathbb{R}^n$ and $w \in \mathbb{R}^{d-n}$, and an open set $V \subseteq \mathbb{R}^n$, and a smooth function $f: V \to \mathbb{R}^{d-n}$, so that the graph $\Gamma(f, V)$ coincides with the surface $S$ near $x_0$. The function $\psi(x) = (x, f(x))$ is a homeomorphism onto its image, and so the set $U = \psi(V)$ is an open neighborhood of $x_0$ in $S$, with a homeomorphism $\varphi = \psi^{-1}: U \to V$. \qed

Example 1.8 (The $\mathbb{R}P^n$). As shown in the previous section, the $n$-sphere $S^n$ is a smooth surface, and so by the previous proposition, $S^n$ is an $n$-dimensional manifold. Let’s look specifically at some standard coordinate charts on it. For each $j \in \{1, \ldots, n+1\}$, consider the (open) half spaces:

$$H_j^+ = \{(x^1, \ldots, x^{n+1}) \in \mathbb{R}^n: \pm x^j > 0\}.$$

Define $U_j^\pm = H_j^+ \cap S^n$, the northern and southern hemispheres. It is easy to check that $U_j^\pm$ is the graph of the function $x_j = \pm f(x^1, \ldots, \hat{x^j}, \ldots, x^{n+1})$, where $f: \mathbb{B}^n \to \mathbb{R}$ is the smooth function $f(u) = \sqrt{1 - |u|^2}$. Here $\mathbb{B}^n = \{u \in \mathbb{R}^n: |u| < 1\}$ is the unit ball, and the notation $(x^1, \ldots, \hat{x^j}, \ldots, x^{n+1})$ means $x^j$ is omitted from the list: i.e. this is the point $(x^1, \ldots, \hat{x^j}, \ldots, x^{n+1}) = (x^1, \ldots, x^{j-1}, x^{j+1}, \ldots, x^{n+1}) \in \mathbb{R}^n$.

Now, any point $x \in S^n$ is contained in at least one of the hemispheres $U_1^+, \ldots, U_{n+1}^\pm$. Thus, for a point in $U_j^\pm$, we have the coordinate chart $\phi_j^\pm: U_j^\pm \to \mathbb{B}^n$ given by

$$\phi_j^\pm(x) = (x^1, \ldots, \hat{x^j}, \ldots, x^{n+1})$$

which is a homeomorphism whose inverse is $y \mapsto (y, \pm f(y))$ (reordered so that the $f(y)$ coordinate is in the $j$th entry).

Now, let’s consider a manifold that is not evidently a subset of a Euclidean space.

Example 1.9 (Real Projective Space). The real projective space $\mathbb{R}P^n$ is defined to be the set of all 1-dimensional subspaces of $\mathbb{R}^{n+1}$; that is, points in $\mathbb{R}P^n$ are lines through the origin in $\mathbb{R}^{n+1}$. It is made into a topological space via the quotient map $\pi: \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{R}P^n$, where $\pi(x) = \text{span} \{x\}$. (This means that a set $U$ in $\mathbb{R}P^n$ is open iff $\pi^{-1}(U)$ is open in $\mathbb{R}^{n+1} \setminus \{0\}$.) Note that $\mathbb{R}P^n$ can also be thought of as the quotient of the $n$-sphere by the equivalence relation $x \sim -x$ (i.e. antipodal points are identified). It is left as an exercise to the topologically-minded reader to prove (or look up the proof) that $\mathbb{R}P^n$ is second-countable and Hausdorff.

For $1 \leq j \leq n+1$, let $V_j \subset \mathbb{R}^{n+1} \setminus \{0\}$ be the set of points where $x^j \neq 0$, and let $U_j = \pi(V_j) \subset \mathbb{R}P^n$. Thus $U_j$ is the set of 1-dimensional subspaces of $\mathbb{R}^{n+1}$ that are not parallel to the $x_j = 0$ plane. The preimage $\pi^{-1}(U_j)$ is easily seen to precisely equal $V_j$, which is open, and so by definition $U_j$ is open in $\mathbb{R}P^n$. Since no 1-dimensional subspaces of $\mathbb{R}^{n+1}$ are parallel to all coordinate planes, it follows that every point in $\mathbb{R}P^n$ is contained in at least one of the $U_j$.

Now, define $\varphi_j: U_j \to \mathbb{R}^n$ by

$$\varphi(\pi(x^1, \ldots, x^{n+1})) = \frac{1}{x^j}(x^1, \ldots, \hat{x^j}, \ldots, x^{n+1}).$$

This map is well-defined since $\pi(x) = \pi(\lambda x)$ for any non-zero scalar $\lambda$. Note that $\varphi: \pi: V_j \to \mathbb{R}^n$ is continuous, and so by the definition of the (quotient) topology on $\mathbb{R}P^n$, it follows that $\varphi_j$ is...
continuous. What’s more, it is a homeomorphism: it is straightforward to verify that its inverse is given by
\[ \varphi_j^{-1}(u^1, \ldots, u^n) = \pi(u^1, \ldots, u^{j-1}, 1, u^{j+1}, \ldots, u^n), \]
which is also evidently continuous. Hence, \( \mathbb{R} \mathbb{P}^n \) is locally Euclidean (with patches in \( \mathbb{R}^n \)), and hence is an \( n \)-dimensional manifold.

Finally, an example for building manifolds out of others.

**Example 1.10 (Product Manifolds).** Let \( M_1, \ldots, M_k \) be manifolds of dimensions \( n_1, \ldots, n_k \). Then \( M_1 \times \cdots \times M_k \) is a manifold of dimension \( n_1 + \cdots + n_k \). Indeed, the product of any finite (or countable) collection of second-countable Hausdorff spaces is second-countable Hausdorff. For the local Euclidean property, if \( (p_1, \ldots, p_k) \in M_1 \times \cdots \times M_k \) is a point in the product, then choose a coordinate chart \( (U_j, \varphi_j) \) near each \( p_j \), and note that \( (U_1 \times \cdots \times U_k, \varphi_1 \times \cdots \times \varphi_k) \) is a coordinate chart near \( (p_1, \ldots, p_k) \).

In particular, taking the circle \( S^1 \) (which is the \( n \)-sphere in the case \( n = 1 \)), the \( n \)-**torus** is the product \( \mathbb{T}^n = S^1 \times \cdots \times S^1 = (S^1)^n \). It is a (fun) \( n \)-dimensional manifold.

### 1.3. Smooth Charts and Atlases

Näively, we might expect that a smooth manifold should be one where the charts \((U, \varphi)\) come with smooth functions \( \varphi \). (This is certainly the case for the hemisphere charts on the \( n \)-sphere given in Example 1.8.) However, this doesn’t really make sense in general: a manifold \( M \) is an abstract topological space, and so we do not know how to define smoothness for a function \( \varphi: U \subseteq M \to \mathbb{R}^n \). In Example 1.8, the set \( M \) was constructed already as a subset of \( \mathbb{R}^{n+1} \), and so the notion of smoothness is present. In general, we want to avoid this case, and work with manifolds abstractly.

In fact, we are going to turn this question around: the point will be to define what it means for a function \( f: M \to \mathbb{R}^n \) to be smooth. Again, the näive choice is to define this locally: taking \( p \in M \) and a chart \((U, \varphi)\) at \( p \), we should insist that the composite function \( f \circ \varphi^{-1}: \varphi(U) \subseteq \mathbb{R}^n \to \mathbb{R}^n \) be smooth. This is, in fact, the definition we will use, but we must be careful. In general, there will be many charts at \( p \), and the definition of smoothness should not depend on which one we use. Thus, for this notion of smoothness to be well-defined, we require a consistency condition.

**Definition 1.11.** Let \( M \) be a manifold, and let \((U, \varphi)\) and \((V, \psi)\) be two charts such that \( U \cap V \neq \emptyset \). The **transition map** from \((U, \varphi)\) to \((V, \psi)\) is the composite map
\[ \psi \circ \varphi^{-1}: \varphi(U \cap V) \to \psi(U \cap V). \]

Note that \( \varphi(U \cap V) \) and \( \psi(U \cap V) \) are open subsets of \( \mathbb{R}^n \), and \( \psi \circ \varphi^{-1} \) is a homeomorphism between them. The two charts are said to be **smoothly compatible** if the transition map \( \psi \circ \varphi^{-1} \) is a diffeomorphism.

A collection of charts whose domains cover all of \( M \) is called an **atlas** \( \mathcal{A} \). If all the charts in \( \mathcal{A} \) are smoothly compatible, we call it a **smooth atlas**.

**Example 1.12.** Following Example 1.8, consider the two charts \((U_1^+, \varphi_1^+)\) and \((U_2^+, \varphi_2^+)\). Note that \( U_1^+ \cap U_2^+ \) is the portion of the circle in the first quadrant, and we have \( \varphi_1^+(U_1^+ \cap U_2^+) = \varphi_2^+(U_1^+ \cap U_2^+) = (0, 1) \), and \( \varphi_1^+(x, y) = y \) while \( \varphi_2^+(x, y) = x \). Then \( (\varphi_2^+)^{-1}(x) = (x, \sqrt{1 - x^2}) \), and so the transition map from \( \varphi_2^+ \) to \( \varphi_1^+ \) is
\[ \varphi_1^+ \circ (\varphi_2^+)^{-1}(x) = \sqrt{1 - x^2}. \]
This is a smooth map on the unit interval; in fact it is a diffeomorphism there. Similar calculations show that all the transition maps in the atlas \( \mathcal{A} = \{ (U^\pm_j, \varphi^\pm_j) : 1 \leq j \leq n + 1 \} \) are diffeomorphisms, making this a smooth atlas on \( \mathbb{S}^{n+1} \).

**Example 1.13.** Following Example [19], consider the atlas of charts \((U_j, \varphi_j)\) on \(\mathbb{R}^n\). Assuming, for convenience, that \(i > j\), we can compute that

\[
\varphi_j \circ \varphi_i^{-1}(u^1, \ldots, u^n) = \frac{1}{u^i} \left( u^1, \ldots, \hat{u}^i, \ldots, u^n \right),
\]

which is a smooth map from \(\varphi_i(U_i \cap U_j)\) onto \(\varphi_j(U_i \cap U_j)\). The inverse is given by a similar formula (with the 1 first and omitted variable second), and so the transition maps are all diffeomorphisms. Hence, this is a smooth atlas for \(\mathbb{R}^n\).

**Example 1.14.** Given smooth atlases on \(M_1, \ldots, M_k\), their products give a smooth atlas on \(M_1 \times \cdots \times M_k\).

We would now like to define a **smooth manifold** to be a manifold in possession of a smooth atlas. But this is not careful enough, for two reasons.

- It is possible for a manifold to possess two or more smooth atlases that are not smoothly compatible with each other. This means that the definition of smooth function on the manifold will again be ill-defined: a given function’s smoothness depends on which atlas one chooses to use. Hence, the smooth atlas itself must be a part of the definition – different smooth atlases on the same manifold define different smooth manifolds.
- On the other hand, it may well be that two distinct smooth atlases give rise to precisely the same class of smooth functions. In this case, we do not wish to consider the two manifold-atlas pairs to be distinct smooth manifolds. We could address this point by defining a smooth manifold to be an equivalence class of manifold-atlas pairs, but there is an easier way, as follows.

### 1.4. Smooth Structures.

**Definition 1.15.** Let \(M\) be a manifold. A smooth atlas \(\mathcal{A}\) on \(M\) is called **complete** or **maximal** if it is not properly contained in any larger smooth atlas. That is, \(\mathcal{A}\) is complete if, given any chart \((U, \varphi)\) that is smoothly compatible with all the charts in \(\mathcal{A}\), it follows that \((U, \varphi) \in \mathcal{A}\).

A complete atlas on \(M\) is called a **smooth structure** on \(M\). A **smooth manifold** is a pair \((M, \mathcal{A})\), where \(M\) is a manifold, and \(\mathcal{A}\) is a smooth structure on \(M\).

The apparent disadvantage of the complete atlas definition of a smooth structure is that such objects are very large, and difficult to describe. For example, the set of all charts smoothly compatible with the \(2(n+1)\) charts in the “hemispheres” atlas of \(\mathbb{S}^n\) in Example [18] includes charts with all possible open sets \(U\) as domains, and then on each one a huge uncountable mess of coordinate chart functions \(\varphi\). Not to worry, though: we can still talk about the smooth structure on the sphere by specifying only those \(2(n+1)\) charts, due to the following.

**Lemma 1.16.** Let \(M\) be a manifold, and let \(\mathcal{A}\) be a smooth atlas on \(M\). Then there is a unique smooth structure \(\overline{\mathcal{A}}\) on \(M\) that contains \(\mathcal{A}\). (It is called the **smooth structure determined by \(\mathcal{A}\)**.)

**Proof.** We define \(\overline{\mathcal{A}}\) to be the set of all charts that are smoothly compatible with every chart in \(\mathcal{A}\). Since \(\mathcal{A}\) is a smooth atlas, \(\mathcal{A} \subseteq \overline{\mathcal{A}}\). In particular, \(\overline{\mathcal{A}}\) is an atlas. We need to show three things.
(1) \( \mathcal{A} \) is a smooth atlas (i.e. all its charts are smoothly compatible).
(2) \( \mathcal{A} \) is a smooth structure (i.e. it is complete).
(3) \( \mathcal{A} \) is the unique smooth structure containing \( \mathcal{A} \).

(1) We must show that if \((U, \varphi)\) and \((V, \psi)\) are two charts in \( \mathcal{A} \) with \( U \cap V \neq \emptyset \), they are smoothly compatible. That is, we must show that \( \psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V) \) is smooth. (We actually need to show it is a diffeomorphism, but in the proof we do this for all pairs of charts, including these two in reverse, showing that the inverse is also smooth.)

Fix a point \( p \in U \cap V \), and let \( x = \varphi(p) \in \varphi(U \cap V) \). Since \( \mathcal{A} \) is an atlas, there is some chart \((W, \vartheta)\) in \( \mathcal{A} \) with \( p \in W \). By the definition of \( \mathcal{A} \), \((U, \varphi)\) and \((V, \psi)\) are both smoothly compatible with \((W, \vartheta)\). Hence, both of the maps \( \vartheta \circ \varphi^{-1} \) and \( \varphi \circ \vartheta \) are smooth on their domains. Thus \( \psi \circ \varphi^{-1} = (\psi \circ \vartheta^{-1}) \circ (\vartheta \circ \varphi^{-1}) \) is smooth on a neighborhood of \( x \). This was for an arbitrary point \( x \in \varphi(U \cap V) \), and smoothness is a local property, so \( \psi \circ \varphi^{-1} \) is indeed smooth on \( \varphi(U \cap V) \).

This shows that the two charts are smoothly compatible, and so \( \mathcal{A} \) is an atlas.

(2) We now show that \( \mathcal{A} \) is a complete smooth atlas. But this is pretty easy: since \( \mathcal{A} \subseteq \mathcal{A} \), if \((U, \varphi)\) is any chart that is smoothly compatible with every chart in \( \mathcal{A} \), then it is automatically smoothly compatible with every chart in \( \mathcal{A} \) – which means, by definition, that \((U, \varphi)\) is in \( \mathcal{A} \).

(3) Let \( \mathcal{B} \) be any smooth structure containing \( \mathcal{A} \). Since \( \mathcal{B} \) is an atlas, all its charts are smoothly compatible, and so in particular every chart in \( \mathcal{B} \) is smoothly compatible with every chart in \( \mathcal{A} \). But that means, by definition, that \( \mathcal{B} \subseteq \mathcal{A} \). But \( \mathcal{B} \) is assumed to be complete, and since every chart in \( \mathcal{A} \) is smoothly compatible, it follows that \( \mathcal{B} = \mathcal{A} \), concluding the proof. □

**Example 1.17.** The Euclidean space \( \mathbb{R}^n \) comes with a standard smooth structure, which is the smooth structure determined by the single coordinate chart \((\mathbb{R}^n, \text{Id}_{\mathbb{R}^n})\). But there are other distinct smooth structures on \( \mathbb{R}^n \) as well. For example when \( n = 1 \), consider the chart \((\mathbb{R}, \psi)\) where \( \psi(x) = x^3 \), which is a homeomorphism \( \mathbb{R} \to \mathbb{R} \). By Lemma 1.16 there is a unique smooth structure on \( \mathbb{R} \) containing this chart. It is not the standard smooth structure; for, if it were, the two charts \((\mathbb{R}, \text{Id}_{\mathbb{R}})\) and \((\mathbb{R}, \psi)\) would be smoothly compatible. But note that \( \text{Id}_{\mathbb{R}} \circ \psi^{-1} \) is the map \( y \mapsto y^{1/3} \), which is not smooth.

Let’s consider now a few more examples of smooth manifolds.

**Example 1.18.** Let \( M \) be a smooth manifold, with complete smooth atlas \( \mathcal{A} \). If \( W \subseteq M \) is any open subset, then \( W \) is also a smooth manifold, with smooth structure determined by the atlas \( \{(U, \varphi) : (U, \varphi) \in \mathcal{A}, U \subseteq W\} \). Note: since \( \mathcal{A} \) is complete, its coordinate charts that happened to be contained in \( W \) already included a covering of \( W \), so the resulting collection of restricted charts is indeed an atlas.

So, in particular, any open subset of \( \mathbb{R}^n \) is a smooth manifold. An example of this form gives us another member of the classical family of Lie groups: \( GL(n) \), the group of all invertible \( n \times n \) matrices. Since it can be described as \( \det^{-1}(\mathbb{R} \setminus \{0\}) \), and \( \det \) is continuous while \( \mathbb{R} \setminus \{0\} \) is open, it follows that \( GL(n) \) is an open subset of \( \mathbb{M}_n \cong \mathbb{R}^{n^2} \). Thus, \( GL(n) \) is a smooth manifold (with the subspace smooth structure).

In fact, \( GL(n) \) is an open dense subset of \( \mathbb{M}_n \). To see this, let \( A \in \mathbb{M}_n \) be any matrix. Let \( p_A(\lambda) = \det(A - \lambda I_n) \) be its characteristic polynomial. This polynomial has \( n \) complex roots (the eigenvalues of \( A \)). Let \( r = \min\{|\lambda| : \lambda \neq 0, p_A(\lambda) = 0\} \). Then \( A - \epsilon I_n \) is invertible for \( 0 < \epsilon < r \). This shows \( GL(n) \) is dense in \( \mathbb{M}_n \).

The next example, the Grassmannian manifold of \( k \)-planes in \( \mathbb{R}^n \), is a generalization of real projective space.
Example 1.19 (Grassmannians). For $0 \leq k \leq n$, let $\text{Gr}(n, k)$ denote the set of all $k$-dimensional subspaces of $\mathbb{R}^n$. (So, for example, $\text{Gr}(n+1, 1) = \mathbb{R}P^n$ as in Example 1.9) Problem 7 on Homework 1 shows how to define the standard smooth structure on the Grassmannian that makes it into a smooth manifold.

In fact, Homework 1 Problem 7 relies on Homework 1 Problem 6, which is essentially the same as Lemma 1.35 (the “Smooth Manifold Charts Lemma”) in [1]. We state this again as a proposition here, since it allows us to dispense with a lot of the separate topological preliminaries when working with a smooth manifold. Also, to be slightly more general, instead of using a fixed Euclidean space $\mathbb{R}^n$, we allow the actual model Euclidean space to vary from point to point (as long as the dimension is fixed), specifying a Hilbert space for each chart. Note that all the usual notions (the norm topology, open sets, continuous and smooth maps) make sense in a finite-dimensional Hilbert space, without imposing the coordinate structure of $\mathbb{R}^n$.

Proposition 1.20. Let $M$ be a nonempty set. Let $J$ be an index set. Let $n$ be a positive integer. For each $j \in J$, suppose
- $U_j$ is a nonempty subset of $M$,
- $H_j$ is a real Hilbert space of dimension $n$, and
- $\varphi_j : U_j \to H_j$ is an injective map whose image is open in $H_j$.

Furthermore, suppose that
- there is a countable subset $I \subset J$ so that $\bigcup_{i \in I} U_i = M$, and
- for any two points $p \neq q$ in $M$, either there is a single $U_j$ with $p, q \in U_j$, or there are disjoint $U_j \cap U_k = \emptyset$ with $p \in U_j$ and $q \in U_k$.

Finally, suppose that each of the transition maps
$$\varphi_k \circ \varphi_j^{-1} : \varphi_j(U_j \cap U_k) \to \varphi_k(U_j \cap U_k), \quad j, k \in J$$
is smooth (as a map from an open subset of $H_j$ to $H_k$). Then there is a unique topology on $M$ for which each $\varphi_j$ is a homeomorphism onto its image, making $M$ a topological manifold. Moreover, if $\lambda_j : H_j \to \mathbb{R}^n$ is a given isometric isomorphism, then the set $\mathcal{A} \equiv \{(U_j, \lambda_j \circ \varphi_j) : j \in J\}$ is a smooth atlas on $M$.

The proof of Proposition 1.20 is left as Problem 6 on Homework 1. We will use this proposition frequently. Moreover, from here on, we will use the terms chart and smooth atlas to refer to the nominally more abstract objects above, where the coordinate patches are in finite-dimensional Hilbert spaces (of fixed dimension) whose representation may vary from chart to chart, rather than necessarily a fixed $\mathbb{R}^n$.

Remark 1.21. We could even loosen the a priori requirement that all the model spaces $H_j$ have the same dimension $n$, only requiring that they be finite dimensional. In this case, it follows from the last requirement (that the transitions maps be smooth) that the dimension of $H_j$ is constant along each connected component of $M$; the proof of this is also part of Problem 6 on Homework 1. But this definition would allow, for example, the disjoint union of $S^1$ and $S^2$ to be considered a smooth manifold, meaning dimension would not be well-defined for disconnected manifolds. We will exclude this possibility.
2. Smooth Maps

Now having defined a smooth manifold (with the idea of giving meaning to smooth functions on said manifolds), we can define smooth maps. Note: the words map and function are usually used interchangeably; here, we will try to be consistent about reserving the word function for a map whose codomain is a Euclidean space \( \mathbb{R}^n \), while map could mean a function between any two manifolds.

**Notation 2.1.** A smooth manifold is a pair \( (M,A) \) where \( A \) is a smooth structure on \( M \). It will be more convenient to just say “let \( M \) be a smooth manifold,” with the given smooth structure understood as part of the given data. In this case, if we must refer to the smooth structure explicitly, we will use the notation \( A_M \).

2.1. Definitions, Basic Properties.

**Definition 2.2.** Let \( M \) and \( N \) be smooth manifolds. A map \( F: M \to N \) is called smooth if, for each \( p \in M \), there is a chart \( (U,\varphi) \in A_M \) with \( p \in U \), and a chart \( (V,\psi) \in A_N \) with \( F(p) \in V \), such that \( F(U) \subseteq V \), and the composite map
\[
\psi \circ F \circ \varphi^{-1} : \varphi(U) \to \psi(V)
\]
is smooth.

In other words: a smooth map is a map that is smooth in local coordinates. This is well-defined: suppose we have two charts \( (U_j,\varphi_j) \) at \( p \) and \( (V_j,\psi_j) \) at \( F(p) \) for \( j = 1,2 \). Since they are each chosen from a smooth atlas, the transition maps between them are smooth. Thus, if \( F \) is smooth (in the above sense) with respect to \( (U_1,\varphi_1) \) and \( (V_1,\psi_1) \), then on \( U_1 \cap U_2 \)
\[
\psi_2 \circ F \circ \varphi_1^{-1} = (\psi_2 \circ \psi_1^{-1}) \circ (\psi_1 \circ F \circ \varphi_1^{-1}) \circ (\varphi_1 \circ \varphi_2^{-1})
\]
is a composition of smooth maps, and is also smooth. This is precisely the reason for the smooth compatibility condition in the definition of smooth atlas to begin with.

There is one slight subtlety in the definition: we require that the chart in the codomain \( (V,\psi) \) satisfy \( F(U) \subseteq V \). This is necessary to even make sense of the composition \( \psi \circ F \) on all of \( U \).

**Proposition 2.3.** Let \( M, N \) be smooth manifolds. A smooth map \( F: M \to N \) is continuous.

**Proof.** Fix a point \( p \in M \). Choose charts \( (U,\varphi) \) at \( p \) and \( (V,\psi) \) at \( F(p) \) so that \( F(U) \subseteq V \) and \( \psi \circ F \circ \varphi^{-1} \) is smooth on \( \varphi(U) \). This composite map is therefore continuous. Thus, on \( U \), we have
\[
F|_U = \psi^{-1} \circ (\psi \circ F \circ \varphi^{-1}) \circ \varphi
\]
is a composition of continuous maps, and hence is continuous. This shows that \( F \) is continuous on a neighborhood of the arbitrary point \( p \in M \); thus \( F \) is continuous.

Using the continuity of \( F \), the following alternate characterization of smoothness is often useful.

**Proposition 2.4.** Let \( M, N \) be smooth manifolds. A map \( F: M \to N \) is smooth if and only if the following holds true. For every \( p \in M \), there is a smooth chart \( (U,\varphi) \) at \( p \) and a smooth chart \( (V,\psi) \) containing \( F(p) \) such that \( U \cap F^{-1}(V) \) is open in \( M \) and
\[
\psi \circ F \circ \varphi^{-1} : \varphi(U \cap F^{-1}(V)) \to \psi(V)
\]
is a smooth map.
The proof is left as an exercise. Note: the condition that \( U \cap F^{-1}(V) \) is open here is a replacement for the choice of \( U, V \) with \( F(U) \subseteq V \) in the definition. Without this explicit requirement, the alternate characterization of smoothness would allow for discontinuous maps.

**Example 2.5.** Let \( F: \mathbb{R} \to \mathbb{R} \) be \( F = 1_{[0, \infty)} \): \( F(x) = 0 \) if \( x < 0 \) and \( F(x) = 1 \) if \( x \geq 1 \). Here we take the usual smooth structure determined by \( (\mathbb{R}, \text{Id}_{\mathbb{R}}) \) for \( \mathbb{R} \) in the domain and codomain. For any \( x \neq 0 \), one can take either \((-\infty, 0)\) or \((0, \infty)\) (equipped with the identity map) as a chart in the domain, and \( \mathbb{R} \) in the codomain, to see that \( F \) is smooth (of course) near \( x \). For \( x = 0 \), take \( U = (-\frac{1}{2}, \frac{1}{2}) \) and \( V = (\frac{1}{2}, \frac{3}{2}) \). Then \( F^{-1}(V) = [0, \infty) \), and so \( U \cap F^{-1}(V) = [0, \frac{1}{2}) \), which is not open in \( \mathbb{R} \). Ignoring this condition, however, and noting that all the coordinate functions are \( \text{Id}_{\mathbb{R}} \), the composition \( \psi \circ F \circ \varphi^{-1} \) on \( \varphi(U \cap F^{-1}(V)) = [0, \frac{1}{2}) \) is just the map \( \psi \circ F \circ \varphi^{-1}(x) = x \), which is smooth (in the sense that it is the restriction of a smooth function on \( \mathbb{R} \) to the set \([0, \frac{1}{2})\)). Hence, the condition \( U \cap F^{-1}(V) \) be open is crucial for smoothness to carry its usual meaning (and imply continuity).

The next lemma follows immediately from the definition of smoothness, and its proof is left as an exercise.

**Lemma 2.6 (Smoothness is Local).** Let \( M, N \) be smooth manifolds. A map \( F: M \to N \) is smooth iff for every \( p \in M \), there is an open neighborhood \( U \subseteq M \) containing \( p \) so that \( F|_U \) is smooth.

**Corollary 2.7 (Smooth Gluing Lemma).** Let \( M, N \) be smooth manifolds. Let \( \{U_j : j \in J\} \) be an open cover of \( M \). For each \( j \), suppose \( F_j: U_j \to N \) is a smooth map. Moreover, suppose that \( F_j|_{U_j \cap U_k} = F_k|_{U_j \cap U_k} \) for all \( j, k \in J \). Then there is a unique smooth map \( F: M \to N \) so that \( F|_{U_j} = F_j \) for each \( j \in J \).

**Proof.** For any \( p \in M \), choose some \( U_j \ni p \), and define \( F(p) = F_j(p) \). This is well-defined by the overlapping consistency condition: if \( p \in U_k \) as well, then \( F_k(p) = F_j(p) = F(p) \). Clearly this is the unique function \( F \) with the desired property, so it remains only to show that \( F \) is smooth. This follows from Lemma 2.6 for any \( p \in M \), choose some \( U_j \ni p \); on this open neighborhood, \( F|_{U_j} = F_j \) is smooth, and hence \( F \) is smooth. \( \square \)

### 2.2. Smooth Functions, and Examples.

In the special case that our target manifold is a Euclidean manifold, we refer to a smooth map \( f: M \to \mathbb{R}^k \) as a smooth function. By definition, this means that, for each \( p \in M \), there is a chart \((U, \varphi)\) at \( p \), and a chart \((V, \psi)\) at \( f(p) \) in \( \mathbb{R}^k \), such that \( f(U) \subseteq V \) and \( \psi \circ f \circ \varphi^{-1} \) is smooth on \( \varphi(U) \). In fact, it suffices to always take \( V = \mathbb{R}^k \) and \( \psi = \text{Id} \), giving the following apparently stronger definition.

**Definition 2.8.** Let \( M \) be a smooth manifold, and let \( k \) be a positive integer. A function \( f: M \to \mathbb{R}^k \) is called smooth if, for each \( p \in M \), there is a chart \((U, \varphi)\) so that the function \( f \circ \varphi^{-1}: \varphi(U) \to \mathbb{R}^k \) is smooth. This composite function is called the **coordinate representation** of \( f \) in the chart \((U, \varphi)\).

If this holds true, then evidently \( f \) is smooth in the sense of Definition 2.2 (taking \((V, \psi) = (\mathbb{R}^k, \text{Id})\)). Conversely, if \( f: M \to \mathbb{R}^k \) is smooth in the sense of Definition 2.2 with some codomain chart \((V, \psi)\) satisfying \( f(U) \subseteq V \), then note that

\[
f \circ \varphi^{-1} = \psi^{-1} \circ (\psi \circ f \circ \varphi^{-1})
\]
is well-defined since \( f(U) \) is contained in \( V \), and is a composition of smooth functions, thus is is smooth. Thus, the two definitions are equivalent. A similar argument shows that if \( M \) is also an open subset of a Euclidean space \( \mathbb{R}^n \), then this notion of smoothness coincides with the usual calculus definition.

In the special case \( k = 1 \), the set of smooth functions \( f: M \to \mathbb{R} \) is denoted \( C^\infty(M) \). In this case, since we can add and multiply the valued \( f(p) \) for any \( p \in M \), \( C^\infty(M) \) has the structure of a commutative algebra. This algebra can be used to recover the topological properties of the manifold; this approach is the beginning of algebraic geometry.

Let us now consider some examples of smooth maps.

**Example 2.9.** Let \( M = \{(x, y) \in \mathbb{R}^2 : x > 0\} \). The function \( f(x, y) = x^2 + y^2 \) is smooth in the classical sense. Let’s consider its coordinate representation in a different chart (other than \( \text{local polar coordinate chart} \)).

**Example 2.10.** It is easy to verify that all of the following maps are smooth on any manifold \( M \):

- Any constant map.
- The identity map \( M \to M \).
- If \( U \subseteq M \) is open, then the inclusion map \( U \hookrightarrow M \) is smooth.

**Example 2.11.** Let \( \iota: \mathbb{S}^n \to \mathbb{R}^{n+1} \) denote the inclusion. Then, using the atlas of hemispheres (cf. Example 1.8) with \( \varphi_j^+ (x^1, \ldots, x^{n+1}) = (x^1, \ldots, \hat{x}^j, \ldots, x^{n+1}) \), we have the coordinate representations of \( \iota, \iota^* \varphi_j^+ (U_j^+) \to \mathbb{R}^{n+1} \) given by

\[
\iota(u^1, \ldots, u^n) = (u^1, \ldots, u^{j-1}, \pm \sqrt{1 - (u^1)^2 - \cdots - (u^n)^2}, u^j, \ldots, u^n)
\]

which is smooth on the domain \( \varphi_j^+ (U_j^+) = \mathbb{B}^n = \{(u^1, \ldots, u^n): (u^1)^2 + \cdots + (u^n)^2 < 1\} \). Thus, \( \iota \) is a smooth map.

**Example 2.12.** Let \( \pi: \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{R}^n \) be the defining projection \( \pi(x) = \text{span}_{\mathbb{R}} \{x\} \). Then, in the coordinate charts \( (U_j, \varphi_j) \) for \( \mathbb{R}^n \) given in Example 1.9 and using the identity coordinates on \( \mathbb{R}^{n+1} \setminus \{0\} \) (restricted to the set \( \{x_j \neq 0\} \) so that \( \pi \) maps them into \( U_j \)), we have

\[
\varphi_j \circ \pi(x^1, \ldots, x^{n+1}) = \frac{1}{x^j}(x^1, \ldots, \hat{x}^j, \ldots, x^{n+1}).
\]

This is a smooth map on its domain, and so \( \pi: \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{R}^n \) is smooth.

We can also build new smooth functions from old ones in the usual ways.

**Proposition 2.13.** Let \( M, N, P \) be smooth manifolds. Let \( k \) be a positive integer.

(a) If \( f, g: M \to \mathbb{R}^k \) are smooth functions, then \( f + g: M \to \mathbb{R}^k \) is a smooth function, and \( f \cdot g: M \to \mathbb{R} \) (the dot product in the range) is a smooth function.

(b) If \( F: M \to N \) and \( G: N \to P \) are smooth, then so is \( G \circ F: M \to P \).
Proof. Part (a) is left as an exercise. For part (b): fix a point \( p \in M \). By smoothness of \( G \), there is a chart \((V, \psi)\) at \( F(p) \) in \( N \), and a chart \((W, \vartheta)\) a \( G(F(p)) \) in \( P \) with \( G(V) \subseteq W \), such that \( \vartheta \circ G \circ \psi^{-1} : \psi(V) \to \vartheta(W) \) is smooth. Since \( F \) is smooth, it is continuous, and so \( F^{-1}(V) \) is an open neighborhood of \( p \) in \( M \). By the completeness of the atlas on \( M \), we may choose a chart \((U, \varphi)\) at \( p \) so that \( U \subseteq F^{-1}(V) \), which implies that \( F(U) \subseteq V \). By the smoothness of \( F \) (and the smooth compatibility of the atlas \( \mathcal{A}_M \)), we have \( \psi \circ F \circ \varphi^{-1} : \varphi(U) \to \psi(V) \) is smooth. Now, \[(G \circ F)(U) = G(F(U)) \subseteq G(V) \subseteq W, \text{ and on } \varphi(U),\] is a composition of smooth maps between Euclidean spaces, and so is smooth. Thus, by definition, \( G \circ F \) is smooth. \(\square\)

Example 2.14. The composition \( \pi \circ \iota : \mathbb{S}^n \to \mathbb{R}P^n \), with \( \iota \) and \( \pi \) from Examples 2.11 and 2.12, is therefore a smooth map. This is the map which sense \( x \in \mathbb{S}^n \) to \( \text{span}_\mathbb{R}\{x\} \in \mathbb{R}P^n \). It is a 2-to-1 covering map, since the preimage of any element \( \pi(x) \in \mathbb{R}P^n \) consists of the two antipodal points \( \{x, -x\} \) in \( \mathbb{S}^n \). This is the usual map by which we visualize \( \mathbb{R}P^n \), as the quotient of \( \mathbb{S}^n \) by identifying antipodal points. We now see that this quotient map is smooth.

As usual, “vector-valued” maps are smooth iff all their components are smooth.

Proposition 2.15. Let \( M_1, \ldots, M_k \) and \( N \) be smooth manifolds. Let \( \pi_j : M_1 \times \cdots \times M_k \to M_j \) be the projection map. Then \( \pi_j \) is smooth for each \( j \). Moreover, if \( F : N \to M_1 \times \cdots \times M_k \) is any map, then \( F \) is smooth if and only if \( F_j = \pi_j \circ F : N \to M_j \) is smooth for \( 1 \leq j \leq k \).

The proof of Proposition 2.15 is left as an exercise.

2.3. Diffeomorphisms. Now that we know what “smooth” means for maps between manifolds, we can define the basic morphism in the category of smooth manifolds.

Definition 2.16. Let \( M \), \( N \) be smooth manifolds. A diffeomorphism \( F : M \to N \) is a smooth map that possesses a smooth inverse. If there exists a diffeomorphism \( M \to N \), we say that \( M \) and \( N \) are diffeomorphic; we write \( M \cong N \).

Example 2.17. (1) The open unit ball \( \mathbb{B}^n \) is diffeomorphic to the whole Euclidean space \( \mathbb{R}^n \), via the map \[ F(x) = \left( \frac{x}{\sqrt{1 - |x|^2}} \right) \] whose inverse is \( F^{-1}(y) = \left( \frac{y}{\sqrt{1 + |y|^2}} \right) \). So \( \mathbb{B}^n \cong \mathbb{R}^n \).

(2) Let \( M \) be a smooth manifold and let \((U, \varphi)\) be a chart in a smooth atlas. Then \( \varphi : U \to \varphi(U) \) is a diffeomorphism (its local coordinate representation in the \( \varphi \)-coordinates is the identity map, in either direction). So: a smooth manifold is locally diffeomorphic to \( \mathbb{R}^n \).

One key immediate property of diffeomorphisms is invariance of dimension (which is extremely hard to prove for homeomorphisms).

Theorem 2.18 (Invariance of Dimension). Let \( M, N \) be smooth manifolds. If \( M \cong N \), then \( \dim M = \dim N \).

(Here \( \dim M = n \) if \( M \) is an \( n \)-dimensional manifold.)
Proof. Let \( F : M \to N \) be a diffeomorphism. For any point \( p \in M \), choose smooth charts \((U, \varphi)\) at \( p \) and \((V, \psi)\) and \( F(p) \) with \( F(U) \subseteq V \). Because \( F \) is a homeomorphism, it is an open map, and so \( F(U) \) is open; by restriction if necessary, wlog we may therefore take \( V = F(U) \). Then \( \psi \circ F \circ \varphi^{-1} : \varphi(U) \to \psi(V) \) is a diffeomorphism between Euclidean spaces, with inverse \( \varphi \circ F^{-1} \circ \psi^{-1} \). But we know (cf. Section [0.5]) that diffeomorphisms can only exist between same-dimensional Euclidean spaces. This concludes the proof. \( \square \)

Here is a list of basic properties of diffeomorphisms that are straightforward to verify.

**Proposition 2.19.**

1. A composition of diffeomorphisms is a diffeomorphism.
2. A finite Cartesian product of diffeomorphisms is a diffeomorphism.
3. A diffeomorphism is a homeomorphism.
4. If \( F : M \to N \) is a diffeomorphism between smooth manifolds, and \( U \subseteq M \) is open, then \( F|_U : U \to F(U) \) is a diffeomorphism.
5. The relation \( M \cong N \) (diffeomorphic) is an equivalence relation on smooth manifolds.

The last point in the proposition shows that we really want to think of a smooth manifold as an equivalence class under the diffeomorphism relation: we don’t want to consider two smooth manifolds to be different if they are, in fact, diffeomorphic.

**Example 2.20.** Consider the alternate smooth structure on \( \mathbb{R} \) from Example [1.17] where the smooth atlas is determined by the global chart \( (\mathbb{R}, \varphi(x) = x^3) \). Denote this smooth manifold by \( \tilde{\mathbb{R}} \). This is not the standard smooth structure on \( \mathbb{R} \). However, consider the map \( F : \mathbb{R} \to \tilde{\mathbb{R}}, \ F(x) = x^{1/3} \).

Then in local (i.e. global) coordinates, with \( (\mathbb{R}, \text{Id}) \) be the chart on \( \mathbb{R} \), we have
\[
\varphi \circ F \circ \text{Id}^{-1}(x) = \varphi(x^3) = x, \quad \text{Id} \circ F^{-1} \circ \varphi^{-1}(y) = F^{-1}(y^{1/3}) = y.
\]

Thus, in local coordinates, \( F \) is the identity map, which is clearly a diffeomorphism. We have therefore shown that \( \mathbb{R} \cong \tilde{\mathbb{R}} \).

So \( \mathbb{R} \) and \( \tilde{\mathbb{R}} \) are “the same” after all (as smooth manifolds, they are diffeomorphic). This leads naturally to the question: are there any smooth structures on the topological manifold \( \mathbb{R} \) that are not diffeomorphic to the standard smooth structure? It turns out the answer is no (and we may get to develop all the tools needed to prove that theorem in this course). A wider question is: given two topological manifolds \( M \) and \( N \) that are homeomorphic, is it possible to find smooth structures on them that are not diffeomorphic? This is a very difficult question in general, and is still a topic of much current research. Here are are few answers in specific cases that are pretty mind-blowing.

- If \( \dim M \leq 3 \), then there exists a unique (up to diffeomorphism) smooth structure on \( M \). [Munkres 1960; Moise 1977]
- If \( n \neq 4 \), then the standard smooth structure on \( \mathbb{R}^n \) is the only one up to diffeomorphism. However, there are uncountably many non-diffeomorphic smooth structures on \( \mathbb{R}^4 \). These are called “fake \( \mathbb{R}^4 \)’s”. [Donaldson and Freedman, 1984].
- There are exactly 15 non-diffeomorphic smooth structures on \( \mathbb{S}^7 \). [Milnor, 1956; Kervaire and Smale, 1963].
- If \( \dim M > 3 \), there exists a compact topological \( n \)-manifold which possesses no smooth structures at all.
2.4. **Partitions of Unity.** Since we must always work locally on a manifold, we want to be able to put local objects together into a global object. The “smooth gluing lemma” Corollary 2.7 is an example. But it is too weak for most purposes: there we must know how to extend our smooth function from one neighborhood to another where they are identical on the (big) intersection; this turns out to be generally impossible to arrange if we didn’t already know of the existence of a global smooth function in the first place. If we only wanted continuity, we only need to arrange agreement on the intersections of closed sets, and this is much easier. But there’s no way to get smoothness this way.

The main tool that we will use throughout this course to overcome this technical problem is called a partition of unity. The first step is to refine our cover of coordinate charts so that they do not overlap too much, and the resulting diffeomorphic patches of $\mathbb{R}^n$ are open balls.

Let $M$ be a manifold, and let $\mathcal{O}$ be an open cover. Another open cover $\mathcal{O}'$ of $M$ is called a refinement of $\mathcal{O}$ if, for any $U \in \mathcal{O}'$, there is some $V \in \mathcal{O}$ with $U \subseteq V$. A cover $\mathcal{O}$ is called locally-finite if, for any $p \in M$, there is a neighborhood of $p$ that only intersects finitely many $U \in \mathcal{O}$. Our first lemma here shows the real reason we need to assume our manifolds are second-countable: this guarantees the existence of locally-finite open covers.

**Lemma 2.21.** Let $M$ be a smooth manifold, and let $\mathcal{O}$ be an open cover of $M$. Then there is an open cover $\mathcal{O}'$ of coordinate charts that refines $\mathcal{O}$, is locally-finite, and each coordinate chart in $\mathcal{O}'$ has compact closure, and is diffeomorphic to $\mathbb{R}^n$ (or $\mathbb{B}^n$).

**Proof.** By second-countability, $M$ is $\sigma$-compact (Exercise): there is a countable collection of compact sets $K_1, K_2, K_3, \ldots$ so that $M = \bigcup_n K_n$. Now, let $V_1 \supset K_1$ be an open set with compact closure. Then $\overline{V_1} \cup K_2$ is compact, and similarly has an open neighborhood $V_2$ with compact closure. Continuing this way, we produce a nested sequence $V_1, V_2, V_3, \ldots$ of open sets, with $\overline{V_n} \subset V_{n+1}$, each with compact closure, such that $K_1 \cup \cdots \cup K_n \subset V_n$. It follows that $M = \bigcup_n V_n$. Then we also have $M = \bigcup_n A_n$ where $A_n$ are the closed annular regions $A_n = \overline{V_n} \setminus V_{n-1}$ (where we take $V_0 = \emptyset$). These are all compact.

Fix $n \geq 1$. For each $p \in A_n$, there is some $W \in \mathcal{O}$ with $p \in U$. There is also some chart $(U, \varphi)$ at $p$, and so $U \cap W$ is an open set containing $p$. Let $\mathcal{B}$ be a ball in $\varphi(U \cap W)$ small enough that $\varphi^{-1}(\mathcal{B}) \subseteq V_{n+1} \setminus V_{n-2}$ (this is possible by the continuity of $\varphi$); then $B_p = \varphi^{-1}(\mathcal{B})$ is a neighborhood of $p$ that is diffeomorphic to a ball (via $\varphi$) and is contained in $U$ and in $V_{n+1} \setminus V_{n-2}$. (We call such neighborhoods coordinate balls.) Now, since $A_n$ is compact, we may choose finitely many $B_p$ of these that cover $A_n$. Take $\mathcal{O}'$ to be this countable collection of open sets for $n \geq 1$, which are all by construction diffeomorphic to balls (and since each element is contained in some $V_{n+1}$, it has compact closure). This gives an open cover of $M$ subordinate to $\mathcal{O}$; we need only show it is locally-finite.

Fix any $p \in M$, and let $B \in \mathcal{O}'$ be a coordinate ball containing $p$, that was introduced at stage $n$ above. Then $B \subseteq V_{n+1}$. Now, if $B' \in \mathcal{O}'$ was introduced at stage $m$, then it is not in $V_{m-2}$, so as long as $m - 2 > n + 1$, $B$ does not intersect $B'$. Thus, the set of elements $B'$ that do intersect the given neighborhood $B$ of $p$ are among those that were introduced at stages $1, 2, \ldots, n + 2$. There are only finitely many of these, and so $\mathcal{O}'$ is locally-finite, as desired. \qed

**Exercise 2.21.1.** The above construction produced a locally-finite open cover $\mathcal{O}'$ that is countable. In fact, in any second-countable space, any locally-finite open cover is countable. Prove this.

Another technical result that will be useful in what follows is the shrinking lemma: any locally-finite open cover can be shrunk a little bit (in the sense of shrinking each open set in the cover) and still remain an open cover.
**Lemma 2.22** (Shrinking Lemma). Let $\mathcal{O}$ be a locally-finite open cover of $M$. Then each $U \in \mathcal{O}$ contains a subset $U'$ with $\overline{U'} \subset U$ such that the collection of $U'$ is also a locally-finite open cover.

**Proof.** As discussed above, the cover $\mathcal{O}$ is countable, so list it $\mathcal{O} = \{U_1, U_2, U_3, \ldots\}$. Consider the set

$$C_1 = U_1 \setminus \bigcup_{j \geq 2} U_j.$$

This set is closed: indeed, it is equal to $\overline{U_1} \setminus \bigcup_{j \geq 2} U_j$, for if $p \in \partial U_1$ then, since $U_1$ is open, $p \notin U_1$, and so $p \in \bigcup_{j \geq 2} U_j$. This shows $\overline{U_1} \setminus \bigcup_{j \geq 2} U_j \subseteq U_1 \setminus \bigcup_{j \geq 2} U_j$, and the reverse containment is immediate.

Thus $C_1$ is a closed set which is closed and contained in $U_1$, and satisfies $M = C_1 \cup U_2 \cup U_3 \cup \cdots$. Choose any open set $U_1'$ with the property that $C_1 \subset U_1' \subset U_1$. Then $\{U_1', U_2, U_3, \ldots\}$ is an open cover of $M$. Now proceed by induction: having produced $U_1', \ldots, U_k'$, let

$$C_{k+1} = U_{k+1} \setminus \bigcup_{j=1}^k \bigcup_{m=1}^n U_j'.$$

By the same argument as above, $C_{k+1}$ is closed, contains $U_2$, and $M = U_1' \cup \cdots \cup U_k' \cup C_{k+1} \cup U_{k+2} \cup U_{k+3} \cup \cdots$. Choose any open set $U_{k+1}'$ which satisfies $C_{k+1} \subset U_{k+1}' \subset U_{k+1}$. We need to show that $\{U_1', U_2', U_3', \ldots\}$ is an open cover; this is where local-finiteness comes in. Fix $p \in M$; then there is a finite collection of $U_j$ containing $p$, and so let $n = \max\{j : p \in U_j\}$. By the above induction argument, $\{U_1', U_2', U_3', \ldots, U_n', U_{n+1}, U_{n+2}, \ldots\}$ is an open cover of $M$, and so $p \in U_1' \cup U_2' \cup \cdots U_n' \cup U_{n+1} \cup U_{n+2} \cup \cdots$. But since $p$ is not in $U_j$ for $j > n$, it follows that $p \in U_1' \cup \cdots \cup U_n'$. This holds for any $p$, and so $\{U_1', U_2', U_3', \ldots\}$ indeed form an open cover. (That it is locally-finite follows from the fact that it is a refinement of a locally-finite open cover.)

This brings us to our local tool for building smooth functions: a (smooth) partition of unity.

**Definition 2.23.** Let $\mathcal{O}$ be an open cover of $M$. A **partition of unity subordinate to $\mathcal{O}$** is a collection of functions $\{\psi_U : M \to \mathbb{R} : U \in \mathcal{O}\}$ with the following properties:

- $\text{supp } \psi_U \subset U$
- $0 \leq \psi_U(p) \leq 1$ for all $p \in M$
- The refinement $\mathcal{O}'$ of $\mathcal{O}$, given by $U \in \mathcal{O}'$ iff $\psi_U$ is not identically $0$, is locally-finite.
- $\sum_{U \in \mathcal{O}} \psi_U(p) = 1$ for each $p \in M$.

Note that the last condition (that the functions sum to 1) makes sense in light of the other conditions: for any fixed $p \in M$, the set of $U \in \mathcal{O}'$ containing $p$ is finite, and so the set of $\psi_U$ for which $\psi_U(p) \neq 0$ is finite; hence the sum is always a finite sum.

The main theorem of this section is the assertion that every smooth manifold admits a smooth partition of unity subordinate to the open cover of its smooth structure. This is an enormously powerful theorem for building global smooth functions out of local ones. Indeed, let $\{(U_j, \varphi_j)\}_{j \in J}$ be the smooth structure of $M$. For each $j$, choose some smooth function $\hat{f}_j : \mathbb{R}^n \to \mathbb{R}$. Then $f_j = \hat{f}_j \circ \varphi_j$ is a smooth function on the smooth manifold $U_j$. We can extend it to a function on all of $M$ by setting $f_j = 0$ on $M \setminus U_j$. This is not generally a smooth function, but that doesn’t matter: we can define the global function $f = \sum_j \psi_j f_j$ for a smooth partition of unity subordinate to the open cover $\{U_j\}_{j \in J}$, and then $f$ will be smooth (because the non-smooth parts of $f_j$ are smashed to 0 by the compact support of $\psi_j$ inside $U_j$).

We will see how to use this technique in many important examples in what follows. First, we need to see why a smooth partition of unity exists. The key is the existence of “smooth bump functions” on $\mathbb{R}^n$. 


Proposition 2.24. Let \( 0 < r < R < \infty \). There exists a smooth function \( h: \mathbb{R}^n \to \mathbb{R} \) with the following properties:

1. \( h(x) = 1 \) for \(|x| \leq r\).
2. \( h(x) = 0 \) for \(|x| \geq R\).
3. \( 0 < h(x) < 1 \) for \( r < |x| < R \).

Proof. First, it suffices to prove the proposition in the case \( n = 1 \): if \( h_1 \) is such a function in the \( n = 1 \) case, then we can set \( h(x) = h_1(|x|) \), which clearly possesses properties (1)–(3). The function \( x \mapsto |x| \) is smooth on \( \mathbb{R}^n \setminus \{0\} \), and so \( h \) (being a composition of a smooth function with \( x \mapsto |x| \)) is as well; it is also smooth at 0 since it is, by construction, \( = 1 \) on a neighborhood of 0.

To prove the existence of such an function \( h_1: \mathbb{R} \to \mathbb{R} \), we first build a smooth version of a cutoff function. Set \( f(t) = e^{-1/t} \mathbb{1}_{t \geq 0} \). Then \( f \) is smooth on \( \mathbb{R} \setminus \{0\} \). In fact, it is also smooth at 0. First, continuity at 0 follows by l'Hôpital's rule since \( \lim_{t \to 0} e^{-1/t} = 0 = f(0) \). Continuing by induction, it is straightforward to verify that, for \( t \neq 0 \) and any positive integer \( k \), \( f^{(k)}(t) = p_k(t)e^{-1/t}/t^{2k} \) for some polynomial \( p_k \). This means, by l'Hôpital’s rule, that \( \lim_{t \to 0} f^{(k)}(t) = 0 \). On the other hand, we also have

\[
\frac{f^{(k)}(0)}{t} = \lim_{t \to 0} \frac{f^{(k-1)}(t) - f^{(k-1)}(0)}{t} = \lim_{t \to 0} \frac{p_{k-1}(t)e^{-1/t}/t^{2(k-1)}}{t} = 0
\]

again by induction. Thus, \( f^{(k)} \) exists and is continuous, for each \( k \), and so \( f \) is smooth, as desired.

Note that \( f(t) = 0 \) for \( t \leq 0 \), while \( f(t) > 0 \) for \( t > 0 \). It is easy to then verify that

\[
h_1(t) = \frac{f(R - |t|)}{f(R - |t|) + f(|t| - r)}
\]

has the desired properties, concluding the proof. \( \square \)

Corollary 2.25. Let \( M \) be a smooth manifold. Let \( K \subset U \subset M \) with \( K \) compact and \( U \) open. Then there is a smooth function \( f: M \to [0,1] \) with \( f \equiv 1 \) on \( K \) and \( f \equiv 0 \) on \( M \setminus U \).

Proof. For each \( p \in K \), choose a chart \((U_p, \varphi_p)\) at \( p \) such that \( U_p \subseteq U \), and \( \varphi_p(p) = 0 \). Let \( R_p > 0 \) be such that the ball \( \mathbb{B}(0, 2R_p) \) is contained in \( \varphi_p(U_p) \). Choose some \( r_p \in (0, R_p) \), and let \( h_p \) be a smooth bump function as in Proposition 2.24 with radii \( r_p \) and \( R_p \). Then \( g_p = h_p \circ \varphi_p \) is a smooth function on \( U_p \). We can extend it to all of \( M \) by setting \( g_p(q) = 0 \) for \( q \in M \setminus \varphi_p^{-1}(\mathbb{B}(0, R_p)) \) (this agrees with the bump function on the preimage of the annulus \( \mathbb{B}(0, 2R_p) \setminus \mathbb{B}(0, R_p) \), and so the function is smooth by the smooth gluing lemma, Corollary 2.7). Hence, for each \( p \) we have a smooth function \( g_p \) which is 0 outside \( U_p \), and hence outside \( U \), and \( g_p \) is strictly positive on \( U'_p = \varphi_p^{-1}(\mathbb{B}(0, R_p)) \).

Now, as \( K \) is compact, we can choose finitely many \( p_1, \ldots, p_k \) so that \( K \subset U'_{p_1} \cup \cdots \cup U'_{p_k} \). Thus, the function \( g = g_{p_1} + \cdots + g_{p_k} \) is a smooth function on \( M \) which is 0 outside \( U \), and which is strictly positive on \( K \). By the compactness of \( K \), this means that there is some \( \delta > 0 \) with \( g(p) \geq \delta \) for all \( p \in K \). Let \( k: \mathbb{R} \to \mathbb{R} \) be a smooth transition function on the interval \([0, \delta]\): \( k(x) = 0 \) for \( x \leq 0 \) and \( k(x) = 1 \) for \( x \geq \delta \), while \( 0 \leq k(x) \leq 1 \) for all \( x \). For example, if \( h_1 \) is a bump function as in Proposition 2.24 with inner radius \( \delta \) and outer radius \( 2\delta \), then we could take

\[
k(x) = \begin{cases} h_1(\delta - x), & x \leq \delta \\ 1, & x \geq \delta. \end{cases}
\]

Then the function \( f = k \circ g \) has the desired properties. \( \square \)
Theorem 2.26. Let $M$ be a smooth manifold, and let $\mathcal{O}$ be an open cover. There is a smooth partition of unity subordinate to $\mathcal{O}$.

Proof. To begin, refine the open cover of smooth charts to a locally finite one $\mathcal{O}$ all of whose elements are coordinate balls with compact closure, cf. Lemma 2.21 and define $\psi_U \equiv 0$ for $U \notin \mathcal{O}$. Now, apply the shrinking lemma to produce a new locally-finite cover $\mathcal{O}'$ refining $\mathcal{O}$ so that for each $U \in \mathcal{O}$ there is a $U' \in \mathcal{O}'$ with $\overline{U'} \subset U$. By our construction, the open sets $U \in \mathcal{O}$ have compact closure, and therefore $\overline{U'}$ is compact. By Corollary 2.25 there is a smooth function $f_U: M \to [0, 1]$ such that $f_U \equiv 1$ on $U'$ and $\text{supp } f_U \subset U$. Define $f = \sum_{U \in \mathcal{O}} f_U$. As the cover is locally-finite, this is a finite sum at each point. For each $p$, there is some neighborhood $V$ so that only finitely many $U \in \mathcal{O}$ intersect $V$; thus, on $V$, $f$ is this fixed finite sum of smooth functions, and so is smooth. Moreover, Since the $U'$ cover $M$, it follows that $f > 0$ on $M$. Thus, we may define $\psi_U = f_U/f$. These are smooth functions, whose supports are subordinate to the locally-finite cover $\mathcal{O}$; they clearly take values in $[0, 1]$, and have been designed so that $\sum_{U \in \mathcal{O}} \psi_U = 1$. □

2.5. Applications of Partitions of Unity.

Proposition 2.27. Let $M$ be a smooth manifold, and let $A, B \subset M$ be disjoint closed sets. There exists a smooth function $f: M \to [0, 1]$ such that $f^{-1}(0) = A$ and $f^{-1}(1) = B$.

Proof. This is an exercise on Homework 2. □

We may talk about smooth functions on any subset of a smooth manifold.

Definition 2.28. Let $M, N$ be smooth manifolds, and let $A \subseteq M$ be any subset. A map $F: A \to N$ is called smooth if, for each $p \in A$, there is an open neighborhood $U_p \subseteq M$ and a smooth map $F_p: U_p \to N$ so that $F|_{A \cap U_p} = F_p|_{A \cap U_p}$.

One might expect the definition of smoothness on $A$ to be the nominally stronger statement that $F$ is the restriction of a smooth function on a neighborhood of all of $A$. As it turns out, these are equivalent for closed sets $A$ and functions taking values in $\mathbb{R}^k$.

Proposition 2.29. Let $A \subseteq M$ be closed, and let $f: A \to \mathbb{R}^k$ be a smooth function in the sense of Definition 2.28. Let $U$ be any open neighborhood of $A$. Then there is a smooth function $\tilde{f}: M \to \mathbb{R}^k$ such that $f|_A = \tilde{f}$ and $\text{supp } \tilde{f} \subseteq U$.

Proof. For each $p \in M$, choose a neighborhood $U_p$ and a local smooth extension $\tilde{f}_p$ of $f$ to $U_p$, as in Definition 2.28, wlog assume $U_p \subseteq U$ (i.e. replace $U_p$ by $U_p \cap U$ if necessary). Then the collection $\{U_p: p \in M\} \cup \{M \setminus A\}$ is an open cover of $M$. Fix a smooth partition of unity $\{\psi_p: p \in M\} \cup \{\psi_0\}$ subordinate to this cover (so, in particular, $\text{supp } \psi_p \subset U_p$ and $\text{supp } \psi_0 \subset M \setminus A$).

For each $p \in A$, the function $\psi_p \tilde{f}_p$ is smooth on $U_p$ and 0 outside the closed subset $\text{supp } \psi_p \subset U_p$; hence, we can extend it to a smooth function on all of $M$ by setting it equal to 0 outside $\text{supp } \psi_p$. Then define

$$\tilde{f} = \sum_{p \in A} \psi_p \tilde{f}_p.$$

By local-finiteness, at each point $q \in M$ there is a neighborhood where this is a fixed finite sum, and thus defined a smooth function on that neighborhood; thus $\tilde{f} \in C^\infty(M, \mathbb{R}^k)$. Note that $\psi_0 = 0$.
on \( A \). Thus, by the partition of unity property, we have for \( q \in A \),
\[
1 = \sum_{p \in M} \psi_p(q) + \psi_0(q) = \sum_{p \in M} \psi_p(q),
\]
and so since \( \tilde{f}_p(q) = f(q) \) for \( q \in A \), we have
\[
\tilde{f}(q) = \sum_{p \in M} \psi_p(q) \tilde{f}_p(q) = \sum_{p \in M} \psi_p(q) f(q) = f(q).
\]
This shows \( \tilde{f} \) is a global smooth extension of \( f \) as desired. We need only show that \( \text{supp} \tilde{f} \subseteq U \).

This is left as an exercise. \( \square \)

**Remark 2.30.** We may use a partition of unity here since the codomain of \( f \) is \( \mathbb{R}^k \) where we can employ vector space operations. If we don’t have this, not only does the above proof not work, but the result is generally false. For example: the identity map \( f : \mathbb{S}^1 \to \mathbb{S}^1 \) is smooth, both in the intrinsic sense of a smooth map between smooth manifolds, and in the above sense of a map being smooth on a closed subset of \( \mathbb{R}^2 \); but this map has no smooth (or even continuous) extension to a map \( \mathbb{R}^2 \to \mathbb{S}^1 \). Indeed, suppose \( f : \mathbb{R}^2 \to \mathbb{S}^1 \) is \( C^1 \) and satisfies \( f(u) = u \) for \( u \in \mathbb{S}^1 \). Then the curve \( \gamma_r(t) = (r \cos t, r \sin t) \) in \( \mathbb{R}^2 \) has a \( C^1 \) image \( f \circ \gamma_r \) in \( \mathbb{S}^1 \). We can compute the winding number \( n(\gamma_r, 0) \) of this curve about 0 in the usual way:
\[
n(\gamma_r, 0) = \oint_{f \circ \gamma_r} \frac{xdy - ydx}{\sqrt{x^2 + y^2}}.
\]
By assumption \( f \circ \gamma_r \) has image contained in \( \mathbb{S}^1 \), so this is the same as
\[
n(\gamma_r, 0) = \oint_{f \circ \gamma_r} xdy - ydx
\]
\[
= \int_0^{2\pi} f_1(r \cos t, r \sin t) \frac{d}{dt} f_2(r \cos t, r \sin t) - f_2(r \cos t, r \sin t) \frac{d}{dt} f_1(r \cos t, r \sin t) dt
\]
\[
= \int_0^{2\pi} f_1(r \cos t, r \sin t) [-\partial_1 f_2(r \cos t, r \sin t) r \sin t + \partial_2 f_2(r \cos t, r \sin t) r \cos t] dt
\]
\[
- \int_0^{2\pi} f_2(r \cos t, r \sin t) [-\partial_1 f_1(r \cos t, r \sin t) r \sin t + \partial_2 f_1(r \cos t, r \sin t) r \cos t] dt.
\]
Because \( f_j \) and \( \partial_i f_j \) are continuous functions for \( i, j \in \{1, 2\} \), this is evidently a continuous function of \( r \). It is also integer valued (as a winding number), so it must be constant. By assumption, when \( r = 1 \), \( n(\gamma_1, 0) = 1 \), and so \( n(\gamma_r, 0) = 1 \). But, by inspection, \( n(\gamma_0, 0) = 0 \), a contradiction.

Another application of partitions of unity that demonstrates their typical application is the existence of a *smooth exhaustion function*. This is a smooth function \( f : M \to \mathbb{R} \) with the property that \( f^{-1}(-\infty, c] \) is compact for each \( c \in \mathbb{R} \). Since \( f \) is defined everywhere, each point in \( M \) is in one (and hence all larger) of these *sublevel sets*. On \( \mathbb{R}^n \), \( f(x) = |x|^2 \) is a smooth exhaustion function; on \( \mathbb{B}^n \), \( f(x) = \frac{1}{1 - |x|^2} \) is one.

For any such function, the collection \( K_n = f^{-1}(-\infty, n] \) forms a nested sequence of compact sets that covers \( M \). So a smooth exhaustion function is a kind of smooth version of such a sequence. (Of course, this only really makes sense for a non-compact \( M \).)

**Proposition 2.31.** Let \( M \) be a smooth manifold. Then there exists a strictly positive smooth exhaustion function \( f : M \to (0, \infty) \).
Proof. Let \( \{V_j\} \) be a locally-finite (thence countable) open cover of \( M \), such that each \( V_j \) has compact closure. Let \( \{\psi_j\} \) be a smooth partition of unity subordinate to this cover. Define \( f = \sum j \psi_j \). For any \( p \in M \), there is a neighborhood on which only finitely many \( \psi_j \) are non-zero, and so \( f \) is smooth on a neighborhood of \( p \), thus \( f \in C^\infty(M) \). Moreover, \( f(p) \geq \sum_j \psi_j(p) = 1 \), so \( f \) is strictly positive.

Now, fix \( c \in \mathbb{R} \), and let \( N > c \) be a positive integer. If \( p \notin \bigcup_{j=1}^N V_j \), since \( \text{supp} \psi_j \subset V_j \), it follows that \( \psi_j(p) = 0 \) for \( 1 \leq j \leq N \). Thus

\[
f(p) = \sum_{j=N+1}^{\infty} j \psi_j(p) \geq N \sum_{j=N+1}^{\infty} \psi_j(p) = N \sum_{j=1}^{\infty} \psi_j(p) = N > c.
\]

That is: if \( p \notin \bigcup_{j=1}^N V_j \) then \( f(p) > c \). The contrapositive statement is: if \( f(p) \leq c \), then \( p \in \bigcup_{j=1}^N V_j \). So the sublevel set \( f^{-1}(-\infty, c] \), which is closed since \( f \) is continuous, is contained in the compact set \( \bigcup_{j=1}^N V_j \), and hence is compact. This shows \( f \) is a smooth exhaustion function. \( \square \)
3. Tangent Vectors

3.1. Tangent Spaces. We typically view element in $\mathbb{R}^n$ with dual meaning: either as points in the metric space, or as vectors in the vector space. In particular, if $f: \mathbb{R}^n \to \mathbb{R}^m$, then the derivative in a direction $v \in \mathbb{R}^n$ at a point $x \in \mathbb{R}^n$ is $D_v f(x) = [Df(x)]v$. Here we view $v$ as a vector in $\mathbb{R}^n$, which we must therefore draw as an arrow anchored at the origin; but geometrically we think of it as anchored at the point $x$. We could be more formal about this.

**Definition 3.1.** The tangent space to $\mathbb{R}^n$ as a point $x \in \mathbb{R}^n$ is the vector space $\{x\} \times \mathbb{R}^n$. We denote it by $T_x \mathbb{R}^n$.

This notation helps us record that the vectors are tangent to the point $x$. This is helpful when we want to talk about vectors tangent to a surface, for example.

**Example 3.2.** The smooth surface $S^n \subset \mathbb{R}^{n+1}$ has a tangent space at each point $x$, consisting of those vectors that are orthogonal to the position vector $x$ itself. That is:

$$T_x S^n = \{(x, v) \in T_x \mathbb{R}^{n+1} : v \perp x\} \subset T_x \mathbb{R}^n,$$

a subspace of the tangent space to $\mathbb{R}^{n+1}$ at $x$.

This notation allows us to conveniently talk about tangent spaces to smooth surfaces imbedded in $\mathbb{R}^n$. We would like to talk about tangent vectors to manifolds that are not (a priori) imbedded in $\mathbb{R}^n$. The presentation above of $T_x S^n$ requires the imbedding: the tangent vectors live in the larger space $\mathbb{R}^{n+1}$ (even if they form an $n$-dimensional subspace thereof). We therefore need another way to talk about such tangent vectors that is intrinsic to the manifold itself.

There are many different (but, in the end, isomorphic) ways to do this. The one we will follow is the most common: having essentially defined the smooth structure on $M$ by selecting which functions will be called smooth, we also define tangent vectors in terms of smooth functions. The key observation is that the space $T_x \mathbb{R}^n$ can be identified with the space of directional derivative differential operators $D_v |_x$ at the point $x$, for $v \in \mathbb{R}^n$. We will therefore define tangent vectors to be such operators. The question is: how can we intrinsically define such operators without already having an explicit $v \in \mathbb{R}^n$ to deal with? The answer lies in the product rule: for any directional derivative $D_v |_x$ at the point $x \in \mathbb{R}^n$ (acting on $C^\infty(\mathbb{R}^n)$), we have the product rule

$$D_v (fg)|_x = f(x) D_v g|_x + g(x) D_v f|_x.$$

It will turn out that this property uniquely defines all directional derivative operators at $x$. We call operators satisfying this property derivations.

**Definition 3.3.** Let $M$ be a smooth manifold, and $p \in M$. The space of derivations $\text{Der}_p(M)$ at $p$ is the space of linear operators $X_p: C^\infty(M) \to \mathbb{R}$ with the property

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

Let us record two key properties of derivations that follow immediately from the definition.

**Lemma 3.4.** Let $M$ be a smooth manifold and $p \in M$, and let $X_p \in \text{Der}_p M$. Then

(a) for any constant function $f \in C^\infty(M)$, $X_p f = 0$, and

(b) if $f, g \in C^\infty(M)$ with $f(p) = g(p) = 0$, then $X_p (fg) = 0$.

**Proof.** For part (a), note that if $f \equiv c$ then $f = c \cdot 1$ where $1$ is the constant function taking value $1$. Thus, by linearity $X_p f = X_p (c \cdot 1) = c X_p 1$, and so it suffices to prove that $X_p 1 = 0$. This follows from the fact that $1 = 1^2$ and the derivation property:

$$X_p 1 = X_p (1^2) = 1(p)X_p 1 + 1(p)X_p 1 = 2X_p 1$$

and the derivation property:

$$X_p (ab) = X_p (a \cdot b) = X_p (a) b + a X_p b$$

and the derivation property:

$$X_p (f + g) = X_p f + X_p g.$$
and this implies that $X_p\mathbb{1} = 0$. Part (b) follows directly from the derivation property:
\[ X_p(fg) = f(p)X_p g + g(p)X_p f = 0 + 0 = 0. \]

\[ \square \]

Elementary as these properties are, they actually allow us to prove that, on $\mathbb{R}^n$, $\text{Der}_x \mathbb{R}^n$ consists exactly of the directional derivative operators $D_v|_x$ for $v \in \mathbb{R}^n$.

**Proposition 3.5.** Fix $p \in \mathbb{R}^n$. The map $\vartheta_p: (p, v) \mapsto D_v|_p$ is an vector space isomorphism $T_p \mathbb{R}^n \to \text{Der}_p \mathbb{R}^n$.

**Proof.** We already noted that directional derivative operators $D_v|_p$ are derivations at $p$, and so the map is well-defined. It is straightforward to verify that it is a linear map. Now, let $v \in \ker \vartheta_p$, so that $D_v f(p) = 0$ for all $f \in C^\infty(\mathbb{R}^n)$. Taking $f(x) = x^j$ for $1 \leq j \leq n$, we have $0 = [D f(p)] v = e^j \cdot v = v^j$, the $j$th component of the vector in the standard basis. This shows that $v = 0$, and so $\vartheta_p$ is one-to-one.

It remains only to show that it is onto. Fix $X_p \in \text{Der}_p \mathbb{R}^n$. Motivated by the previous argument, define the vector $v$ by taking its $j$th component in the standard basis to be $X_p(x \mapsto x^j)$. We will show that $X_p = \vartheta_p(p, v) = D_v|_x$. By Taylor’s Theorem $0.9$, we have
\[ f(x) = f(p) + \sum_{i=1}^n \partial_i f(p)(x^i - p^i) + \frac{1}{2} \sum_{i,j=1}^n (x^i - p^i)(x^j - p^j) \int_0^1 (1-t)\partial_i \partial_j f(p + t(x-p)) \, dt. \]

From Proposition $0.11$, the functions $x \mapsto \int_0^1 (1-t)\partial_i \partial_j f(p + t(x-p)) \, dt$ are smooth. Thus, setting $g^i(x) = x^i - p^i$ and $h^j(p) = (x^j - p^j) \int_0^1 (1-t)\partial_i \partial_j f(p + t(x-p)) \, dt$, the functions $g^i, h^j$ are smooth and satisfy $g^i(p) = h^j(p) = 0$ for $1 \leq i, j \leq n$, and we have
\[ f(x) = f(p) + \sum_{i=1}^n \partial_i f(p)(x^i - p^i) + \frac{1}{2} \sum_{i,j=1}^n g^i(x) h^j(x). \]

By Lemma $3.4$ it follows that
\[ X_p f = \sum_{i=1}^n X_p \left( \partial_i f(p)(x^i - p^i) \right) = \sum_{i=1}^n \partial_i f(p) v^i = D_v f|_p \]
as desired. \[ \square \]

We thus have an intrinsic realization of the geometric tangent space $T_p \mathbb{R}^n$: we can view it as the space of derivations $\text{Der}_p \mathbb{R}^n$. So, for example, instead of talking about the vector $[1, 0, -2]^T$ tangent to $p \in \mathbb{R}^3$, we could instead think of this vector as the first-order differential operator $(\partial_1 - 2\partial_3)|_p$ acting on $C^\infty(\mathbb{R}^3)$. With this description in mind, we are prompted to define tangent spaces on manifolds accordingly.

**Definition 3.6.** Let $M$ be a smooth manifold, and $p \in M$. The tangent space $T_p M$ to $M$ at $p$ is defined to be the space $\text{Der}_p M$ of derivations at $p$. That is, the tangent space is the set of linear operators $X_p: C^\infty(M) \to \mathbb{R}$ satisfying $X_p(fg) = f(p)X_p g + g(p)X_p f$ for all $f, g \in C^\infty(M)$.

There is an important technical point to understand here: while the derivations formally act on smooth functions defined on all of $M$, they really only depend on the behavior of the functions in an arbitrarily small neighborhood of $p$. 


**Lemma 3.7.** Let \( X_p \in \text{Der}_p M \), let \( f, g \in C^\infty(M) \), and let \( U \) be any open neighborhood of \( p \) in \( M \). If \( f|_U = g|_U \), then \( X_p(f) = X_p(g) \).

**Proof.** Let \( h = f - g \), so \( h \in C^\infty(M) \) satisfies \( h = 0 \) on \( U \). The support \( \text{supp} h \) is a closed set disjoint from \( U \). Now, let \( \psi \) be a bump function that is identically equal to 1 on \( \text{supp} h \), and satisfies \( \psi(p) = 0 \). Then \( \psi h \) is identically equal to \( h \): if \( h(p) \neq 0 \), then \( (\psi h)(p) = 1 \cdot h(p) = h(p) \), and otherwise \( (\psi h)(p) = 0 = h(p) \). But then, by Lemma 3.4, since \( \psi(p) = h(p) = 0 \), we have \( X_p(h) = X_p(\psi h) = 0 \). By linearity of \( X_p \), this means \( 0 = X_p(h) = X_p(f - g) = X_p(f) - X_p(g) \), so \( X_p(f) = X_p(g) \) as desired. \( \square \)

By virtue of Proposition 3.5, when \( M = \mathbb{R}^n \), \( \text{Der}_p M \) is an \( n \)-dimensional vector space. A priori, it may not be clear that this space is finite-dimensional (or non-zero) on a general manifold. We will shortly see that it has the same dimension at every point \( p \), equal to \( \dim M \). Indeed, this will follow from the fact that \( M \) looks like \( \mathbb{R}^n \) in some neighborhood of \( p \). To understand how this works, we first need to understand the generalized version of the total derivative.

### 3.2. The Differential / Tangent Map

For smooth functions \( F: U \subseteq \mathbb{R}^n \to \mathbb{R}^m \), we have the total derivative \( DF(p): \mathbb{R}^n \to \mathbb{R}^m \), the best linear approximation of \( F \) near \( p \). In fact, in our new language of anchoring vectors at base points, it is clear that we should think of \( DF(p) \) as a linear map from \( T_p \mathbb{R}^n \) to \( T_{F(p)} \mathbb{R}^m \). That is, after all, how we draw the associated vectors. It also makes a great deal of sense in terms of the chain rule: if we have \( F: \mathbb{R}^n \to \mathbb{R}^m \) and \( G: \mathbb{R}^m \to \mathbb{R}^k \), then we have \( D(G \circ F)(p): T_p \mathbb{R}^n \to T_{G(F(p))} \mathbb{R}^k \) satisfies \( D(G \circ F)(p) = DG(F(p)) \circ DF(p) \), where \( DF(p): T_p \mathbb{R}^n \to T_{F(p)} \mathbb{R}^m \), and so we see that the evaluation point \( F(p) \) in \( DG \) is necessary for the composition to even make sense, so that \( DG(F(p)) : T_{F(p)} \mathbb{R}^m \to T_{G(F(p))} \mathbb{R}^k \).

Now, by Proposition 3.5, we may view \( T_p \mathbb{R}^n \) as \( \text{Der}_p \mathbb{R}^n \) via the isomorphism \( \partial_p \). Thus, we can view \( DF(p): T_p \mathbb{R}^n \to T_{F(p)} \mathbb{R}^m \) as a linear map \( \text{Der}_p \mathbb{R}^n \to \text{Der}_{F(p)} \mathbb{R}^m \); i.e. let us define

\[
\mathcal{D}_p F = \partial F \circ DF \circ \partial_p^{-1}.
\]

This conjugated total derivative is called the **differential** of \( F \) at \( p \). That is, for any \( X_p \in \text{Der}_p \mathbb{R}^n \), \( \mathcal{D}_p F(X_p) \) is a derivation in \( \text{Der}_p \mathbb{R}^m \). So

\[
\mathcal{D}_p F(X_p) = \partial F(DF(\partial_p^{-1}(X_p))).
\]

Let \( \partial_p^{-1}(X_p) = (p, v) \). Then \( DF(p)(\partial_p^{-1}(X_p)) = [DF(p)](p, v) = (F(p), [DF(p)]v) \in T_{F(p)} \mathbb{R}^m \). The action of \( \partial F \) is to convert a vector into the directional derivative operator in the direction of that vector, and so

\[
\mathcal{D}_p F(X_p) = D(DF(v)|_{F(p)}).
\]

In other words, if \( f \in C^\infty(\mathbb{R}^m) \), then \( \mathcal{D}_p F(X_p) \) is the derivation in \( \text{Der}_{F(p)} \mathbb{R}^m \) with action

\[
\mathcal{D}_p F(X_p)(f) = D([DF(v)](f)|_{F(p)}) = [DF(F(p))][DF(v)]v.
\]

Now, employing the chain rule, this means that

\[
\mathcal{D}_p F(X_p)(f) = [DF(F)(f)]v.
\]

On the other hand, naïvely, given an element \( X_p \in \text{Der}_p M \), how could we make it act on \( C^\infty(\mathbb{R}^m) \)? Well, for any \( f \in C^\infty(\mathbb{R}^m) \), the push-forward function \( f \circ F \) is in \( C^\infty(\mathbb{R}^n) \), which is the domain of the operator \( X_p \). Since we have \( X_p = \partial_p \), we have

\[
X_p(f \circ F) = D_v(f \circ F)|_p = [DF(F)(f)]v.
\]

Comparing (3.2) and (3.3) leads to the following immediate, but deep, observation.
Thus $d \in U$.

We will prove (d), leaving the other parts to the reader. Let $p \colon F \mapsto \text{map}$ see it denoted as true (since derivations in $\text{Der}$ language) is nothing more than the (pointwise) push-forward by $F$: the natural composition map needed to fit the pieces together (thanks to the chain rule).

We can thus discuss the differential of a smooth map between manifolds.

**Definition 3.9.** Let $M, N$ be smooth manifolds, and let $F \colon M \to N$ be smooth. For $p \in M$, the differential $dF_p$ is the map $T_p M \to T_{F(p)} N$ defined by

$$dF_p(X_p)(f) = X_p(f \circ F), \quad X_p \in T_p M, \quad f \in C^\infty(N).$$

Note that, if $f, g \in C^\infty(N)$, then setting $Y_{F(p)} = dF_p(X_p)$ we have

$$Y_{F(p)}((f \cdot g)) = X_p((f \cdot g) \circ F) = X_p((f \circ F) \cdot (g \circ F)) = (f \circ F)(p)X_p(g \circ F) + (g \circ F)(p)X_p(f \circ F) = f(F(p))Y_{F(p)}(g) + g(F(p))Y_{F(p)}(f).$$

That is: $Y_{F(p)}$ is indeed in $\text{Der}_{F(p)} N$, and so $dF_p$ is well-defined.

The collection of names/notations for this map is essentially uncountable. Some authors call it the differential, others the total derivative, still others the unimaginative tangent map. You might see it denoted as $dF_p = DF(p) = F'(p) = T_F p$, or even as $dF = DF = F' = TF$ with the $p$ suppressed. Another popular notation is $F_*$, which highlights the fact that the map (in this language) is nothing more than the (pointwise) push-forward from $\text{Der}_p M$ to $\text{Der}_{F(p)} N$, via the map $F$. We will try to be consistent with the notation $dF_p$, and the name differential.

Here are some basic properties of differentials of smooth maps than can be readily verified from the definition.

**Proposition 3.10.** Let $M, N, P$ be smooth manifolds, with $F \colon M \to N$ and $G \colon N \to P$ smooth maps. Let $p \in M$.

1. $dF_p : T_p M \to T_{F(p)} N$ is a linear map.
2. $d(G \circ F)_p = dG_{F(p)} \circ dF_p$.
3. $d(\text{id}_M)_p = \text{id}_{T_p M}$.
4. If $F$ is a diffeomorphism, then $dF_p$ is a linear isomorphism, and $(dF_p)^{-1} = d(F^{-1})_{F(p)}$.

**Proof.** We will prove (d), leaving the other parts to the reader. Let $q = F(p)$; then we have $dF_p : T_p M \to T_q N$ and $d(F^{-1})_q : T_q N \to T_p M$. We compose them: for $X_p \in T_p M, Y_q \in T_q N, f \in C^\infty(M), \text{and } g \in C^\infty(N)$,

$$[dF_p(X_p)(f)](g) = [dF_p(Y_q)(g)](f) = X_p((f \circ F) \circ F)(g) = Y_q((f \cdot g) \circ F) = Y_q(g).$$

Thus $dF_p \circ d(F^{-1})_q = \text{id}_{T_p M}$ and $dF_p \circ d(F^{-1})_q = \text{id}_{T_q N}$. □

From our (pre-derivation) definition of the tangent space $T_p M$ to $\mathbb{R}^n$, it’s clear that if we restrict to an open subset $U \ni p$, $T_p U = T_p \mathbb{R}^n$. In terms of derivations, this equality cannot be strictly true (since derivations in $\text{Der}_p \mathbb{R}^n$ act on a difference space than those in $\text{Der}_p U$). Nevertheless, the differential allows us to identify them in a natural way, on a general smooth manifold. After all, Lemma 3.7 shows that derivations act locally.
**Proposition 3.11.** Let $M$ be a smooth manifold, and let $U \subseteq M$ be open. Denote by $\iota : U \hookrightarrow M$ the inclusion map (which is smooth). For any $p \in U$, the differential $d\iota_p : T_p U \to T_p M$ is a linear isomorphism.

**Proof.** Suppose $X_p \in T_p U$ is in $\ker (d\iota_p)$. Fix any $f \in C^\infty (U)$, and some open neighborhood $B$ of $p$ such that $\overline{B} \subset U$. By Proposition 2.29 there is a smooth extension $\tilde{f} \in C^\infty (M)$ so that $\tilde{f}|_U = f|_U$. Then $f$ and $\tilde{f}|_U$ are smooth functions in $C^\infty (U)$ that agree on the open neighborhood $B$ of $p$, and so by Lemma 3.7 $X_p(f) = X_p(\tilde{f}|_U)$. But $\tilde{f}|_U = \tilde{f} \circ \iota$, and so

$$X_p(f) = X_p(\tilde{f} \circ \iota) = d\iota_p(X_p)(\tilde{f}) = 0.$$ 

This shows that $X_p = 0$. Thus, $d\iota_p$ is injective.

For surjectivity, let $Y_p \in T_p M$. We define $X_p \in T_p U$ as follows: as above, fix some open neighborhood $B$ of $p$ with $\overline{B} \subset U$, and for $f \in C^\infty (U)$, define $X_p(f) = Y_p(\tilde{f})$ where $\tilde{f}$ is any smooth function on $M$ which agrees with $f$ on $B$. This is well-defined by Lemma 3.7 and it is easy to verify that $X_p$ is a derivation. We claim that $d\iota_p(X_p) = Y_p$. Indeed, for any $g \in C^\infty (M)$, we have

$$d\iota_p(X_p)(g) = X_p(g \circ \iota).$$

By definition, $X_p(g \circ \iota) = Y_p(h)$ for any smooth extension of $g \circ \iota|_U$ to $M$. The function $g$ is such an extension, and so $d\iota_p(X_p)(g) = Y_p(g)$, as desired. Thus, $d\iota_p$ is surjective, completing the proof. \[\square\]

We therefore canonically identify $T_p U \cong T_p M$, keeping in mind that the derivations in $T_p U$ act on functions in the larger space $C^\infty (M)$ by acting on any smooth extension from a function in a neighborhood of $p$ in $U$.

**Remark 3.12.** One way to avoid this slight complication is to define derivations a little differently: in light of Lemma 3.7, we may think of the domain of a derivation $X_p$ not as the space $C^\infty (M)$, but as the space of equivalences classes of smooth functions that agree on any neighborhood of $p$. Such an equivalence class is called a germ, and the space of germs at $p$ is usually denoted $C^\infty_p (M)$. Then we might define the tangent space to be the set of derivations of $C^\infty_p (M)$. This makes the local nature of the tangent space clearer, and it greatly simplifies the proof of Proposition 3.11 (since $C^\infty_p (M) = C^\infty_p (U)$). In particular, this means we don’t need to use smooth extensions, whose existence depends on bump functions. If we were interested in analytic manifolds, we would have no choice but to use the germ definition, since analytic functions are too rigid to allow the kinds of restriction/extension arguments above. But we will stick with the present formalism, since the topic is already highly abstract, and we don’t want to complicate it more than necessary.

With this precise statement that the tangent space is local, combined with the local Euclidean nature of the manifold, we can quickly prove that $T_p M$ is an $n$-dimensional vector space at each $p \in M$ (where $n = \dim M$).

**Corollary 3.13.** If $M$ is an $n$-dimensional smooth manifold, then for each $p \in M$, $T_p M$ is an $n$-dimensional vector space.

**Proof.** Fix $p \in M$, and let $(U, \varphi)$ be a chart at $p$. Then $\varphi : U \to \hat{U}$ is a diffeomorphism onto an open set $\hat{U} \subseteq \mathbb{R}^n$. By Proposition 3.10(d), it follows that $d\varphi_p$ is an isomorphism $T_p U \to T_{\varphi(p)} \hat{U}$. By Proposition 3.11, we have $T_p U \cong T_p M$, and $T_{\varphi(p)} \hat{U} \cong T_{\varphi(p)} \mathbb{R}^n$. By Proposition 3.5, the latter is isomorphic to $\mathbb{R}^n$, concluding the proof. \[\square\]
One more general useful identification is the tangent space to a product, which can be thought of as the direct sum of the tangent spaces.

**Proposition 3.14.** Let $M_1, \ldots, M_k$ be smooth manifolds, and let $M = M_1 \times \cdots \times M_k$. Denote by $\pi_j: M \to M_j$ the projection map for $1 \leq j \leq k$. Fix a point $p = (p_1, \ldots, p_k) \in M$. Then the map

$$d(\pi_1)_p \oplus \cdots \oplus d(\pi_k)_p: T_pM \to T_{p_1}M_1 \oplus \cdots \oplus T_{p_k}M_k$$

is a linear isomorphism.

The proof is left as an exercise.

### 3.3 Local Coordinates

Tangent vectors are now viewed as abstract derivations. However, we can make these concrete in local coordinates. Fix a point $p \in M$, and let $(U, \varphi)$ be a chart at $p$. Denote by $\hat{p} = \varphi(p)$ and $\hat{U} = \varphi(U)$. Let $i: U \hookrightarrow M$ and $\hat{i}: \hat{U} \hookrightarrow \mathbb{R}^n$ to be the inclusion maps. Then, following the proof of Corollary 3.13,

$$di_{\hat{p}} \circ d\varphi_p \circ (di_p)^{-1}: T_pM \to T_{\hat{p}}\mathbb{R}^n$$

is a linear isomorphism. Note: we will usually ignore the maps $di$ and $di$ and consider them to be the identity: that is, we use the isomorphism property of Proposition 3.11 to identify the tangent space of an open subset of a manifold with the tangent space of the manifold. Thus, we will (somewhat informally, but consistently) think of $d\varphi_p$ as an isomorphism from $T_pM$ onto $T_{\hat{p}}\mathbb{R}^n$.

Now, $T_{\hat{p}}\mathbb{R}^n$ has a canonical basis: starting with the canonical basis $\{e^1, \ldots, e^n\}$ of $\mathbb{R}^n$, the images under the isomorphism $\partial_{\hat{p}}$ are the derivations $D_{e^j}\big|_{\hat{p}} = \frac{\partial}{\partial x^j}\big|_{\hat{p}}$ for $1 \leq j \leq n$. Since $d\varphi_p$ is an isomorphism, the preimages of these derivations form a basis for $T_pM$. Here is another standard abuse of notation: we denote these preimages $\frac{\partial}{\partial x^j}\big|_{\hat{p}}$:

$$\left. \frac{\partial}{\partial x^j}\right|_{\hat{p}} \equiv (d\varphi_p)^{-1}\left(\left. \frac{\partial}{\partial x^j}\right|_{\hat{p}}\right) = d(\varphi^{-1})_{\hat{p}}\left(\left. \frac{\partial}{\partial x^j}\right|_{\hat{p}}\right).$$

So $\frac{\partial}{\partial x^j}\big|_{\hat{p}}$ is a derivation acting on $C^\infty(M)$ (or $C^\infty(U)$). How does it act? Untwisting the definition, given $f \in C^\infty(U)$, we write it in local coordinates as $\hat{f} = f \circ \varphi^{-1}: \hat{U} \to \mathbb{R}$. Then by the definition of the differential, we have

$$\left. \frac{\partial}{\partial x^j}\right|_{\hat{p}}(f) = d(\varphi^{-1})_{\hat{p}}\left(\left. \frac{\partial}{\partial x^j}\right|_{\hat{p}}(f) = \left. \frac{\partial}{\partial x^j}\right|_{\hat{p}}(f \circ \varphi^{-1}) = \frac{\partial \hat{f}}{\partial x^j}(\hat{p}).$$

In other words: a canonical basis for $T_pU \cong T_pM$ is given by the $n$ derivations whose actions are to take the partial derivatives of smooth functions in the chart coordinates. The vectors $\frac{\partial}{\partial x^1}\big|_{\hat{p}}, \ldots, \frac{\partial}{\partial x^n}\big|_{\hat{p}}$ are called the **coordinate vectors at** $p$; it is important to note that they depend on the chart $\varphi$ of choice.

**Example 3.15.** Let $M = \mathbb{R}^3$, and consider the chart $(U, \varphi)$ where $U = \{(x, y, z) \in \mathbb{R}^3: x > 0\}$, and $\varphi: U \to (0, \infty) \times (-\frac{\pi}{2}, \frac{\pi}{2}) \times (0, \pi)$ is the spherical polar map whose inverse is $\varphi^{-1}(\rho, \theta, \phi) = (\rho \cos \theta \sin \phi, \rho \sin \theta \sin \phi, \rho \cos \phi)$. Fix a point $p = (x_0, y_0, z_0) \in U$, with spherical polar coordinates $\hat{p} = \varphi(p) = (\rho_0, \theta_0, \phi_0)$. The 3 coordinate vectors at this point are the derivations

$$\left. \frac{\partial}{\partial \rho}\right|_{\hat{p}}, \left. \frac{\partial}{\partial \theta}\right|_{\hat{p}}, \left. \frac{\partial}{\partial \phi}\right|_{\hat{p}} \in T_{\hat{p}}\hat{U}. \quad (3.4)$$
These are just the images, under the isomorphism \( \partial_{\rho} \), of the standard basis vectors \( e^1, e^2, e^3 \) in \((\rho, \theta, \phi)\)-space. The more interesting question of expressing them as vectors in terms of the original \((x, y, z)\)-coordinates is discussed below, in Example 3.18.

Thus, every vector \( X_p \in T_p M \) can be expressed uniquely as a linear combination

\[
X_p = \sum_{j=1}^{n} X^j_p \frac{\partial}{\partial x^j} |_p, \quad X^j_p \in \mathbb{R}, \ 1 \leq j \leq n.
\]

How do we compute the components \( X^j_p \) for a given \( X_p \)? Well, for any \( f \in C^\infty(U) \),

\[
X_p(f) = \sum_{j=1}^{n} X^j_p \frac{\partial f}{\partial x^j} |_p = \sum_{j=1}^{n} X^j_p \frac{\partial f}{\partial x^j}(\hat{\rho}).
\]

So, in particular, taking \( \hat{f}(x) = x^j \) yields simply \( X_p(f) = X^j_p \). Thus, we compute the components by evaluating \( X_p \) on the \( n \) smooth functions \( f \) for which \( \hat{f} = f \circ \varphi^{-1}(x) = x^j, \ 1 \leq j \leq n \).

What are these functions? They are \( f(p) = \varphi^j(p) \), where \( \varphi = (\varphi^1, \ldots, \varphi^n) \) (the \( n \) component functions of \( \varphi: U \rightarrow \mathbb{R}^n \)). So \( X^j_p = X_p(\varphi^j); \) this is often (abusively) written as \( X_p(x^j) \), writing \( \varphi(p) = (x^1(p), \ldots, x^n(p)) \).

**Example 3.16.** Following Example 3.15, consider the derivation \( \frac{\partial}{\partial z}|_p \in T_p U \) (where \( U \) is the \( x > 0 \) open half-plane). We want to express \( \frac{\partial}{\partial z}|_p \) as a linear combination of the coordinate vectors (3.4).

To do this, we need expressions for the coordinates \( (\varphi^1, \varphi^2, \varphi^3) = (\rho, \theta, \phi) \) (rather than the inverse \( \varphi^{-1}(\rho, \theta, \phi) = (x, y, z) \)). Once we have these, we then have

\[
\frac{\partial}{\partial z} = \frac{\partial \rho}{\partial z} \frac{\partial}{\partial \rho} |_p + \frac{\partial \theta}{\partial z} \frac{\partial}{\partial \theta} |_p + \frac{\partial \phi}{\partial z} \frac{\partial}{\partial \phi} |_p.
\]

(For example, we know \( \rho = \sqrt{x^2 + y^2 + z^2} \) so \( \frac{\partial \rho}{\partial z} = \frac{\partial \rho}{\partial \phi} \) is \( \cos \phi \).)

Now that we can express vectors in local coordinates, we can also write the differential in local coordinates. Let \( F: M^m \rightarrow N^n \) be a smooth map between smooth manifolds. Let \((U, \varphi)\) be a chart at \( p \), and let \((V, \psi)\) be a chart at \( F(p) \). Then we have a coordinate representation of \( F \):

\[
\hat{F} = \psi \circ F \circ \varphi^{-1}: \varphi(U \cap F^{-1}(V)) \rightarrow \psi(V).
\]

As above, let \( \hat{p} = \varphi(p) \). We denote the components of \( \varphi \) as \((x^1, \ldots, x^m)\) and the components of \( \psi \) as \((y^1, \ldots, y^n)\).

**Proposition 3.17.** In terms of the coordinate bases \( \left\{ \frac{\partial}{\partial y^j} |_{F(p)} \right\}_{1 \leq j \leq n} \) for \( T_{F(p)} N \), we have

\[
dF_p \left( \frac{\partial}{\partial x^j} |_p \right) = \sum_{k=1}^{n} \frac{\partial \hat{F}^k}{\partial x^j}(\hat{p}) \frac{\partial}{\partial y^j} |_{F(p)} , \quad 1 \leq j \leq m.
\]

In other words: in local coordinates, the matrix of \( dF_p \) is precisely the Jacobian matrix of the coordinate representation \( \hat{F} \) of the function. The differential has been cooked up to be a coordinate-independent version of the Jacobian matrix.
Proof. By definition,
\[ \frac{\partial}{\partial x^j} \bigg|_p = d(\varphi^{-1})_{\hat{p}} \left( \frac{\partial}{\partial x^j} \bigg|_\hat{p} \right), \]
and so
\[ dF_p \left( \frac{\partial}{\partial x^j} \bigg|_p \right) = d(F \circ \varphi^{-1})_{\hat{p}} \left( \frac{\partial}{\partial x^j} \bigg|_\hat{p} \right). \]
Now, since \( F \circ \varphi^{-1} = \psi^{-1} \circ \hat{F} \), we therefore have
\[ d(F \circ \varphi^{-1})_{\hat{p}} \left( \frac{\partial}{\partial x^j} \bigg|_\hat{p} \right) = d(\psi^{-1})_{\hat{F}(\hat{p})} \left( \frac{\partial}{\partial x^j} \bigg|_{\hat{F}(\hat{p})} \right) = d(\psi^{-1})_{\hat{F}(\hat{p})} \left( \frac{\partial}{\partial \hat{x}^j} \bigg|_{\hat{p}} \right) d(\hat{F})_{\hat{p}} \left( \frac{\partial}{\partial \hat{x}^j} \bigg|_{\hat{p}} \right). \]
Now, the inside expression is the differential \( d\hat{F}_{\hat{p}} \) of a map between Euclidean spaces, acting on a derivation on Euclidean space. By definition (3.1), this is the total derivative \( D\hat{F}(\hat{p}) \) acting on the vector corresponding (via \( \vartheta_{\hat{p}} \)) to \( \frac{\partial}{\partial x^j} \bigg|_\hat{p} \) (which is the unit basis vector \( e^j \)) expressed in terms of derivations in the coordinates \( \frac{\partial}{\partial \hat{x}^j} \bigg|_{\hat{F}(\hat{p})} \). The matrix of the total derivative is just the Jacobian matrix, and so we have
\[ d\hat{F}_{\hat{p}} \left( \frac{\partial}{\partial \hat{x}^j} \bigg|_{\hat{p}} \right) = \vartheta_{\hat{F}(\hat{p})} \left( \frac{\partial}{\partial \hat{x}^j} \bigg|_{\hat{p}} \right) \frac{\partial \hat{F}_k}{\partial x^j}(\hat{p}) e_k = \sum_{k=1}^n \frac{\partial \hat{F}_k}{\partial x^j}(\hat{p}) \frac{\partial}{\partial y^k} \bigg|_{\hat{F}(\hat{p})} \frac{\partial}{\partial \hat{x}^j} \bigg|_{\hat{p}}. \]
Using linearity and the fact (by definition) that
\[ d(\psi^{-1})_{\hat{F}(\hat{p})} \left( \frac{\partial}{\partial \hat{x}^j} \bigg|_{\hat{F}(\hat{p})} \right) = \frac{\partial}{\partial \hat{x}^j} \bigg|_{\hat{F}(\hat{p})}, \]
proves the result. \( \square \)

As a final exercise in our discussion of tangent vectors in local coordinates, let us consider a change of coordinates. That is: consider two charts \((U, \varphi)\) and \((V, \psi)\) at \( p \in M \), and as usual denote the components of \( \varphi \) as \((x^1, \ldots, x^n)\) and the components of \( \psi \) as \((y^1, \ldots, y^n)\). Then the transition map \( \psi \circ \varphi^{-1} \) is the map taking \((x^1, \ldots, x^n)\) to \((y^1, \ldots, y^n)\). This is a special case of the above discussion, where we take the identity function \( F = \text{Id}_M : M \rightarrow M \) and express it in these local coordinates: \( \psi \circ \varphi^{-1} = \psi \circ \text{Id}_M \circ \varphi^{-1} = \text{Id}_M \). Hence, from Proposition 3.17, we have
\[ \frac{\partial}{\partial x^j} \bigg|_p = d(\text{Id}_M)_p \left( \frac{\partial}{\partial x^j} \bigg|_p \right) = \sum_{k=1}^n \left( \frac{\partial (\text{Id}_M)^k}{\partial x^j}(\hat{p}) \right) \frac{\partial}{\partial y^k} \bigg|_{\text{Id}_M(\hat{p})} \frac{\partial}{\partial \hat{x}^j} \bigg|_{\text{Id}_M(\hat{p})} \right. \]
\[ = \sum_{k=1}^n \left( \frac{\partial y^k}{\partial \hat{x}^j}(\hat{p}) \right) \frac{\partial}{\partial \hat{x}^j} \bigg|_{\hat{p}} \frac{\partial}{\partial \hat{x}^j} \bigg|_{\hat{p}}. \] (3.5)
Again, this is just another statement of the chain rule.

Example 3.18. Continuing Examples 3.15 and 3.16, our manifold is \( M = \mathbb{R}^3 \), and we have some base point \( p \) in the \( x > 0 \) half-plane. We have two sets of coordinates: \( \varphi = (\rho, \theta, \phi) \) and \( \psi = (x, y, z) \). From (3.5), we can express the coordinate vectors in the spherical coordinates in terms
of the Euclidean ones, via

\[
\frac{\partial}{\partial \rho} = \frac{\partial x}{\partial \rho} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \rho} \frac{\partial}{\partial y} + \frac{\partial z}{\partial \rho} \frac{\partial}{\partial z}
\]

\[
\frac{\partial}{\partial \theta} = \frac{\partial x}{\partial \theta} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta} \frac{\partial}{\partial y} + \frac{\partial z}{\partial \theta} \frac{\partial}{\partial z}
\]

\[
\frac{\partial}{\partial \phi} = \frac{\partial x}{\partial \phi} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \phi} \frac{\partial}{\partial y} + \frac{\partial z}{\partial \phi} \frac{\partial}{\partial z}
\]

where \( \hat{p} = \varphi(p) = (\rho_0, \theta_0, \phi_0) \) is the base point’s spherical coordinates. In this case, we have the explicit coordinate transformation \( (x, y, z) = \psi \circ \varphi^{-1}(\rho, \theta, \phi) = (\rho \cos \theta \sin \phi, \rho \sin \theta \sin \phi, \rho \cos \phi) \), and so we can explicitly write down the Jacobian yielding

\[
\begin{align*}
\frac{\partial}{\partial \rho} & = \cos \theta_0 \sin \phi_0 \frac{\partial}{\partial x} + \sin \theta_0 \sin \phi_0 \frac{\partial}{\partial y} + \cos \phi_0 \frac{\partial}{\partial z} \\
\frac{\partial}{\partial \theta} & = -\rho_0 \sin \theta_0 \sin \phi_0 \frac{\partial}{\partial x} + \rho_0 \cos \theta_0 \sin \phi_0 \frac{\partial}{\partial y} \\
\frac{\partial}{\partial \phi} & = \rho_0 \cos \theta_0 \cos \phi_0 \frac{\partial}{\partial x} + \rho_0 \sin \theta_0 \cos \phi_0 \frac{\partial}{\partial y} - \rho_0 \sin \phi_0 \frac{\partial}{\partial z}
\end{align*}
\]

Note: by transforming these derivations first via \((d\psi^{-1})_{\psi(p)}\) into partial derivative operators in the \((x, y, z)\) coordinates, and then into vectors in \(T_{\psi(p)}\mathbb{R}^3 \cong \{\psi(p)\} \times \mathbb{R}^3\) via the isomorphism \(\psi_{\psi(p)}\), we get the three vectors

\[
e_\rho = \begin{bmatrix} \cos \theta_0 \sin \phi_0 \\ \sin \theta_0 \sin \phi_0 \\ \cos \phi_0 \end{bmatrix}, \quad e_\theta = \begin{bmatrix} -\rho_0 \sin \theta_0 \sin \phi_0 \\ \rho_0 \cos \theta_0 \sin \phi_0 \\ 0 \end{bmatrix}, \quad e_\phi = \begin{bmatrix} \rho_0 \cos \theta_0 \cos \phi_0 \\ \rho_0 \sin \theta_0 \cos \phi_0 \\ -\rho_0 \sin \phi_0 \end{bmatrix}.
\]

These vectors are routinely used by physicists and engineers. They are the standard basis vectors in spherical coordinates (at the point \(p\) whose spherical coordinates are \((\rho_0, \theta_0, \phi_0)\)).

### 3.4. Velocity Vectors of Curves

Let \(M\) be a smooth manifold. A smooth curve in \(M\) (perhaps more accurately a smooth parametrized curve) is a smooth map \(\alpha: (a, b) \to M\) for some open interval \((a, b) \subseteq \mathbb{R}\). Fix \(t_0 \in (a, b)\), and let \((U, \varphi)\) be a local chart at \(p = \alpha(t_0)\) in \(M\); then we can express the curve in local coordinates as \(\hat{\alpha}(t) = \varphi \circ \alpha(t)\). The velocity vector to the curve \(\hat{\alpha}(t) = (\hat{\alpha}^1(t), \ldots, \hat{\alpha}^n(t))\) is the vector in \(\mathbb{R}^n\)

\[
\left( \frac{d\hat{\alpha}^1}{dt}(t_0), \ldots, \frac{d\hat{\alpha}^n}{dt}(t_0) \right) = \sum_{j=1}^n \frac{d\hat{\alpha}^j}{dt}(t_0)e^j.
\]

Realizing the tangent space \(T_{\hat{\alpha}(t_0)}\hat{U}\) as \(\text{Der}_{\hat{\alpha}(t_0)}(\mathbb{R}^n)\), the velocity vector is

\[
\sum_{j=1}^n \frac{d\hat{\alpha}^j}{dt}(t_0) \frac{\partial}{\partial x^j} \bigg|_{\hat{\alpha}(t_0)} \quad (3.6)
\]
whose action on $\hat{f} \in C^\infty(\mathbb{R}^n)$ is (by the chain rule)

$$\sum_{j=1}^n \frac{d\hat{\alpha}^j}{dt}(t_0) \frac{\partial}{\partial x^j} \bigg| \hat{\alpha}(t_0) (\hat{f}) = \frac{d}{dt} (\hat{f} \circ \hat{\alpha}) \bigg|_{t_0}.$$

Now, $\hat{\alpha} = \varphi \circ \alpha$, and if we take $f \in C^\infty(M)$, restricted to $U$, and let $\hat{f} = f \circ \varphi^{-1}$ as usual, then this means the velocity vector to the local coordinate representation $\hat{\alpha}$ of $\alpha$ at $t_0$ acts as a derivation on $C^\infty(M)$ by

$$\hat{\alpha}(t_0)(f) \equiv \frac{d}{dt} (f \circ \alpha) \bigg|_{t_0}.$$

This is our definition of the velocity vector to the curve $\alpha$ at time $t_0$, which is manifestly coordinate independent. In fact, noting that $\alpha: (a, b) \to M$ is a smooth map between manifolds, we can interpret the velocity vector as

$$\hat{\alpha}(t_0)(f) \equiv \frac{d}{dt} (f \circ \alpha) \bigg|_{t_0} = d\alpha_{t_0} \left( \left. \frac{d}{dt} \right|_{t_0} \right)(f).$$

Here we write $\left. \frac{d}{dt} \right|_{t_0}$ instead of $\left. \frac{\partial}{\partial t} \right|_{t_0}$, as is common when there is only one variable; it is the standard basis vector for $T_{t_0}(a, b)$. So, to summarize:

**Definition 3.19.** Let $(a, b) \subseteq \mathbb{R}$ be a nonempty open interval, let $M$ be a smooth manifold, and let $\alpha: (a, b) \to M$ be a smooth curve. For any $t_0 \in (a, b)$, the velocity vector of $\alpha$ at time $t_0$ is defined to be the derivation

$$\hat{\alpha}(t_0) \equiv d\alpha_{t_0} \left( \left. \frac{d}{dt} \right|_{t_0} \right) \in T_{\alpha(t_0)} M.$$

In local coordinates, it is given by the usual expression (3.6).

In fact, we can use velocity vectors of curves to naturally characterize all tangent vectors. First note that all tangent vectors are velocity vectors to (lots of) curves.

**Lemma 3.20.** Let $M$ be a smooth manifold, $p \in M$, and $X_p \in T_p M$. There is some smooth curve $\alpha: (-\epsilon, \epsilon) \to M$ for some $\epsilon > 0$ with $\alpha(0) = p$ so that $\hat{\alpha}(0) = X_p$.

**Proof.** Let $(U, \varphi)$ be a chart at $p$, and with $\varphi = (x^1, \ldots, x^n)$, write $X_p$ in coordinates $X_p = \sum_{j=1}^n X^j_p \frac{\partial}{\partial x^j}\big|_p$. Because $\hat{U} = \varphi(U)$ is open and $\hat{p} = \varphi(p) \in \hat{U}$, there is some $\epsilon > 0$ so that the straight curve $\hat{\alpha}(t) = t(X^1_p, \ldots, X^n_p)$ for $|t| < \epsilon$ is contained in $\hat{U}$. Define $\alpha(t) = \varphi^{-1}(t(X^1_p, \ldots, X^n_p))$. The above computations (cf. e.velocity.local) show that $\hat{\alpha}(0) = X_p$, as desired. \qed

What’s more: velocity vectors of curves transform naturally under smooth maps.

**Lemma 3.21.** Let $M, N$ be smooth manifolds, and let $F: M \to N$ be a smooth map. If $\alpha: (a, b) \to M$ is a smooth curve in $M$, then $\beta = F \circ \alpha: (a, b) \to N$ is a smooth curve in $N$, and for any $t_0 \in (a, b)$,

$$\dot{\beta}(t_0) = dF_{\alpha(t_0)}(\dot{\alpha}(t_0)).$$
Proof. This is just the chain rule for differentials:
\[ \dot{\beta}(t_0) = d\beta_{t_0} \left( \frac{d}{dt} \big|_{t_0} \right) = d(F \circ \alpha)_{t_0} \left( \frac{d}{dt} \big|_{t_0} \right) = dF_{\alpha(t_0)} \left( d\alpha_{t_0} \left( \frac{d}{dt} \big|_{t_0} \right) \right) = dF_{\alpha(t_0)}(\dot{\alpha}(t_0)). \]

Lemmas [3.20] and [3.21] give a method for computing differentials that is often the most effective in practice.

**Corollary 3.22.** Let \( M, N \) be smooth manifolds, and \( F: M \to N \) a smooth map. Fix \( p \in M \), and \( X_p \in T_pM \). Then
\[ dF_p(X_p) = (F \circ \alpha)'(0) \]
for any smooth curve \( \alpha: (-\epsilon, \epsilon) \to M \) such that \( \alpha(0) = p \) and \( \dot{\alpha}(0) = X_p \).

(Here we have used the “prime” notation instead of the “dot” notation, since it is hard to put a dot over top the composition. We will generally use these two interchangeably for velocity vectors of curves.) If \( F \) is presented in a form other than an explicit coordinate representation, Corollary [3.22] is often the best way to actually calculate the differential \( dF_p \). We will use this frequently in what comes.

**Remark 3.23.** Corollary [3.22] shows us yet another equivalent construction of the tangent space: we can define \( T_pM \) to consist of equivalence classes of smooth curves \( \alpha: (-\epsilon, \epsilon) \to M \) with \( \alpha(0) = p \). We want the equivalence relation to somehow say that all elements should have the same \( \dot{\alpha}(0) \); not having defined the tangent space yet, we instead follow the above discussion and say that two curves \( \alpha, \beta \) are equivalent if, for every smooth function \( f: M \to \mathbb{R} \), the two real-valued curves \( f \circ \alpha \) and \( f \circ \beta \) have the same velocity vector at \( t = 0 \): i.e. \( (f \circ \alpha)'(0) = (f \circ \beta)'(0) \). Thus, we think of tangent vectors based at \( p \) as “infinitesimal curves” through \( p \). The preceding corollary then shows that there is a bijection from this construction to the derivations we’ve chosen to use. What’s more, as Lemma [3.21] hints at, in this construction of \( T_pM \), the differential \( dF_p \) is simply the map which sends a vector \([\alpha]\) (the equivalence class of some curve \( \alpha \)) to \([F \circ \alpha]\); the chain rule then shows this is well-defined. This is a very nice geometric way to picture the tangent space; it has the significant disadvantage that there is no obvious vector space structure (indeed, the best way to see that \( T_pM \) is a vector space is to use the bijection to our tangent space to import it).

### 3.5. The Tangent Bundle

We now have a good understanding of the tangent space at a point, \( T_pM \). We can then put all of them together to form the **tangent bundle**.

**Definition 3.24.** Let \( M \) be a smooth manifold. The **tangent bundle** \( TM \) is the disjoint union
\[ TM = \bigsqcup_{p \in M} T_pM. \]
It carries a natural projection map \( \pi: TM \to M \), defined by \( \pi(X_p) = p \) for \( X_p \in T_pM \).

For example, as we can identify \( T_p\mathbb{R}^n \cong \{p\} \times \mathbb{R}^n \), we have \( T\mathbb{R}^n \cong \bigsqcup_{p \in \mathbb{R}^n} \{p\} \times \mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n \). In some sense, this should work for any manifold, since \( T_pM \cong \mathbb{R}^n \) for every \( p \). However, this is purely set-theoretic; there is more interesting structure we can put on \( TM \), and it typically will not amount to just \( M \times \mathbb{R}^n \). Indeed, \( TM \) can be (naturally) given the structure of a smooth manifold of dimension \( 2n \).
To make $TM$ into a smooth manifold, we refer to Proposition 1.20. we can define the topology and smooth structure all in one shot. We need to define the charts. To that end, begin by selecting some covering set of charts $(U, \varphi)$ of $M$. For each such $U$, let $\tilde{U} = \pi^{-1}(U)$: the collection of all vectors $X_p$ anchored at all points $p \in U$. Now, let $\varphi(p) = (x^1(p), \ldots, x^n(p))$ for $p \in U$. Then any element of $\pi^{-1}(U)$ can be written uniquely in the form

$$X_p = \sum_{j=1}^{n} X^j_p \frac{\partial}{\partial x^j} \bigg|_p.$$ 

We therefore define a coordinate chart $\tilde{\varphi}: \tilde{U} \to \mathbb{R}^n \times \mathbb{R}^n$ by

$$\tilde{\varphi} \left( \sum_{j=1}^{n} X^j_p \frac{\partial}{\partial x^j} \right) = (x^1(p), \ldots, x^n(p), X^1_p, \ldots, X^n_p) = (\varphi(p), X^1_p, \ldots, X^n_p).$$

(3.7)

Notice that the image $\tilde{\varphi}(\tilde{U}) = \varphi(U) \times \mathbb{R}^n$ is an open subset of $\mathbb{R}^n \times \mathbb{R}^n$, and the map is a bijection here – its inverse is given by

$$\tilde{\varphi}^{-1}(x, v^1, \ldots, v^n) = \sum_{j=1}^{n} v^j \frac{\partial}{\partial x^j} \bigg|_{\varphi^{-1}(x)}, \quad x \in \varphi(U).$$

To apply Proposition 1.20, we need to verify the last three conditions for the charts $(\tilde{U}, \tilde{\varphi})$ for $TM$. First, by selecting a countable set of the $(U, \varphi)$ that cover $M$, the corresponding countable collection of $(\tilde{U}, \tilde{\varphi})$ cover $TM$. Moreover, if $X_p$ and $Y_q$ are distinct elements of $TM$, then either $p = q$ and so $X_p$ and $Y_q = Y_p$ both lie in a single chart $(\tilde{U}, \tilde{\varphi})$ (for any $U$ containing $p$), or $p \neq q$ in which case (by the Hausdorff assumption on $M$) there are disjoint charts $U, V$ on $M$ with $p \in U$ and $q \in V$, in which case $\tilde{U}$ and $\tilde{V}$ are disjoint charts containing $X_p$ and $Y_q$.

Thus, it remains only to show that the transition maps are smooth. Let $(U, \varphi)$ and $(V, \psi)$ be charts on $M$ with $U \cap V \neq \emptyset$, and let $(\tilde{U}, \tilde{\varphi})$ and $(\tilde{V}, \tilde{\psi})$ be the corresponding charts on $TM$. Then $\tilde{\varphi}(U \cap V) = \varphi(U \cap V) \times \mathbb{R}^n$ and $\tilde{\psi}(U \cap V) = \psi(U \cap V) \times \mathbb{R}^n$ are open subsets of $\mathbb{R}^n \times \mathbb{R}^n$. We can compute the transition map

$$\tilde{\psi} \circ \tilde{\varphi}^{-1}: \varphi(U \cap V) \times \mathbb{R}^n \to \psi(U \cap V) \times \mathbb{R}^n$$

explicitly. Denote the components of $\varphi$ as $x = (x^1, \ldots, x^n)$ and the components of $\psi$ as $y = (y^1, \ldots, y^n)$. Then for any point $p \in U \cap V$, and any $v \in \mathbb{R}^n$,

$$\tilde{\varphi}^{-1}(x, v) = \sum_{j=1}^{n} v^j \frac{\partial}{\partial x^j} \bigg|_{p} = \sum_{j,k=1}^{n} v^j \frac{\partial y^k}{\partial x^j} \bigg|_{x} \frac{\partial}{\partial y^k} \bigg|_{p}$$

from (3.5). Hence

$$\tilde{\psi} \circ \tilde{\varphi}^{-1}(x, v) = \left( y(p), \sum_{j=1}^{n} v^j \frac{\partial y^1}{\partial x^j}(x), \ldots, \sum_{j=1}^{n} v^j \frac{\partial y^n}{\partial x^j}(x) \right).$$

This is evidently a smooth function of $(x, v)$. Thus, we have satisfied all the conditions of Proposition 3.7 and so $TM$ is a smooth manifold. What’s more: in the chart $(\tilde{U}, \tilde{\varphi})$, we have $\tilde{\pi} = \pi \circ \tilde{\varphi}^{-1}$ has the action $\tilde{\pi}(x, v) = x$, which is smooth.

Let us summarize this discussion as follows.
Proposition 3.25. Let $M$ be a smooth manifold. The charts in (3.1) define a smooth atlas that makes $TM$ into a smooth manifold, and the projection map $\pi : TM \to M$ is a smooth map.

As noted, while $TM$ is locally isomorphic to $M \times \mathbb{R}^n$, this is typically not the case globally. (We will see some examples a little later on the course.) If $TM \cong M \times \mathbb{R}^n$, we say the tangent bundle is trivial. One case where the tangent bundle is trivial is when the manifold itself has global coordinates.

Proposition 3.26. Let $M$ be a smooth manifold, and suppose there is a single chart covering $M$ (i.e. there exists a diffeomorphism $\varphi : M \to \hat{U}$ for some open subset $\hat{U} \subseteq \mathbb{R}^n$). Then $TM$ is diffeomorphic to $M \times \mathbb{R}^n$.

Proof. In the discussion preceding Proposition 3.25 we showed that for any chart $(U, \varphi)$, the bundle chart $(\hat{U}, \hat{\varphi})$ is a diffeomorphism onto $\varphi(U) \times \mathbb{R}^n$. Since $\hat{M} = \pi^{-1}(M) = TM$, and since $\varphi(M) = \hat{U}$, this proves the proposition. □

Remark 3.27. It can happen that $TM$ is trivial even if there is no global chart for $M$ (indeed, this is true for all Lie groups, which we will study in the following two quarters). But typically $TM$ is non-trivial. For example, it is known (due to deep work of Adams) that $T\mathbb{S}^n$ is trivial iff $n \in \{1, 3, 7\}$. The tangent bundle $T\mathbb{S}^2$ is not diffeomorphic to $\mathbb{S}^2 \times \mathbb{R}^2$ (we will be able to prove this, at least in a restricted form, in a few more pages). What 4-manifold is it? Think of it this way, for general $n$: at each point $p \in \mathbb{S}^n$, we can identify $T_p\mathbb{S}^n \cong \mathbb{R}^n$. We can then further identify this with $\mathbb{S}^n \setminus \{-p\}$, via stereographic projection from the point $p$. Gluing these together, then, we have

$$T\mathbb{S}^n \cong \mathbb{S}^n \times \mathbb{S}^n \setminus \{(p, -p) : p \in \mathbb{S}^n\}.$$ 

It is the product of two spheres with the anti-diagonal removed. When $n = 1$, this is a torus with one complete (twisted) circle “unzipped”, which just gives the cylinder $\mathbb{S}^1 \times \mathbb{R}$ as expected. For $n = 2$, this unzipping does not work, and the bundle is not trivial.

Now, suppose $F : M \to N$ is a smooth map. We have studied its differential $dF_p : T_pM \to T_{F(p)}N$ pointwise. Now that we know how to glue the tangent spaces together to form a manifold, we can glue these maps together as well to form a smooth map between tangent bundles.

Proposition 3.28. Let $M^n$, $N^n$ be smooth manifolds, and let $F : M \to N$ be a smooth map. Define a map $dF : TM \to TN$ as follows: for any $p \in M$ and any $X_p \in T_pM$, $dF(X_p) \equiv dF_p(X_p) \in T_{F(p)}N \subseteq TN$. Then $dF$ is a smooth map (called the global differential of $F$).

Proof. Fix a chart $(U, \varphi)$ in $M$, and look at the action of $dF$ on the neighborhood $\tilde{U} \subset TM$: in local coordinates, we have

$$\tilde{\varphi} \circ dF \circ \tilde{\varphi}^{-1}(x^1, \ldots, x^m, v^1, \ldots, v^m) = \left( F^1(x), \ldots, F^n(x), \sum_{j=1}^{m} \partial_1 F^1(x) v^j, \ldots, \sum_{j=1}^{m} \partial_1 F^n(x) v^j \right).$$

This is smooth (since $F$ is smooth). □

We end this section by stating properties of the global differential, which follow immediately from the same properties of the pointwise differential (cf. Proposition 3.10).

Proposition 3.29. Let $M$, $N$, $P$ be smooth manifolds, with $F : M \to N$ and $G : N \to P$ smooth maps.

(a) $d(Id_M) = Id_{TM}$. 
(b) \( d(G \circ F) = dG \circ dF \).

(c) If \( F \) is a diffeomorphism, then \( dF : TM \to TN \) is a diffeomorphism, and \( d(F^{-1}) = (dF)^{-1} \).
4. Vector Fields

4.1. Definitions and Examples. Let $M$ be a smooth manifold. A vector field on $M$ is simply a choice of a tangent vector at each point. In other words, it is a map $X: M \to TM$, with the property that $X(p) \in T_p M$ for each $p \in M$. This can be written as follows: let $\pi: TM \to M$ be the projection $\pi(X_p) = p$ for any $X_p \in T_p M$. Then a vector field is a function $X: M \to TM$ with the property that

$$\pi \circ X = \text{Id}_M.$$ 

That is: $X$ is a right-inverse to the projection $\pi$. Such a map is generally called a section, so a vector field is a section of the tangent bundle. While $X: M \to TM$ is not valued in $\mathbb{R}^n$, we can still talk about it roughly in those terms since, for each $p$, the value $X(p) = X_p$ of $X$ is in $T_p M$ which is a vector space. In particular, we define as usual the support $\text{supp} X$ of a vector field to be the closure of the set $\{p \in M : X_p \neq 0\}$.

We will largely be concerned with smooth vector fields (in the sense of being a smooth map between the two smooth manifolds $M$ and $TM$). There is a simple characterization of smoothness here (in local coordinates).

**Lemma 4.1.** Let $X$ be a vector field on a smooth manifold $M$; so $X$ is a function $p \mapsto X_p \in T_p M$. Then $X$ is smooth if and only if for any coordinate chart $(U, \phi)$ on $M$ with $\phi = (x^1, \ldots, x^n)$, in the coordinate representation

$$X_p = \sum_{j=1}^n X^j_p \frac{\partial}{\partial x^j} \bigg|_p$$

the functions $p \mapsto X^j_p$ are smooth functions $U \to \mathbb{R}$.

**Proof.** In the natural coordinate chart $(\tilde{U}, \tilde{\phi})$ on $TM$, the representation of the vector field at any point $p \in U$ is given by

$$\tilde{\phi} \left( \sum_{j=1}^n X^j_p \frac{\partial}{\partial x^j} \bigg|_p \right) = (x^1(p), \ldots, x^n(p), \dot{x}^1, \ldots, \dot{x}^n).$$

Thus, if we write $X$ in local coordinates $(U, \phi)$ on $M$ and $(\tilde{U}, \tilde{\phi})$ on $TM$, this gives hat$X = \tilde{\phi} \circ X \circ \phi^{-1}$ is the map $\hat{U} = \phi(U) \to \mathbb{R}^{2n}$ given by

$$\hat{X}(x^1, \ldots, x^n) = (x^1, \ldots, x^n, \dot{x}^1(x), \ldots, \dot{x}^n(x))$$

where $\dot{x}^j(x) = X^j_{\phi^{-1}(x)}$. By definition, of smoothness, $X^j_p$ is smooth iff $\hat{X}^j_p$ is smooth (for every chart), and the map $\hat{X}$ between open subsets of Euclidean space is smooth iff its components are smooth. Thus, $X$ is smooth iff $\hat{X}$ is smooth iff $\hat{X}^j$ are smooth for $1 \leq j \leq n$ if $p \mapsto X^j_p$ are smooth for $1 \leq j \leq n$, as claimed. \hfill $\Box$

**Notation 4.2.** We denote the set of smooth sections of $TM$, i.e. smooth vector fields on $M$, as $\mathcal{X}(M)$. A vector field that is not in $\mathcal{X}(M)$ is often called a rough vector field.

**Example 4.3.** Any vector field in $\mathcal{X}(\mathbb{R}^n)$ has a representation of the form $X = \sum_{j=1}^n X^j \frac{\partial}{\partial x^j}$, for functions $X^1, \ldots, X^n \in C^\infty(\mathbb{R}^n)$. These are the kinds of vector fields discussed in vector calculus. With the identification $X \sim (X^1, \ldots, X^n)$, one can think of such a vector field as a transformation $\mathbb{R}^n \to \mathbb{R}^n$. This is the wrong picture in general, since it means abandoning the structure of having the vector $X_p$ anchored at the point $p$. 

An important specific example is the Euler field \( E \in \mathcal{X}(\mathbb{R}^n) \), given by
\[
E_x = x^1 \frac{\partial}{\partial x^1} + \cdots + x^n \frac{\partial}{\partial x^n}.
\]
It points radially outward, with magnitude equal to the distance of the point from the origin. Thought of as a transformation, it is just the identity map; this bears no relation to its behavior. We will come back to this example later through the chapter.

**Example 4.4.** Realize \( S^1 \) as the unit circle in \( \mathbb{C} \): \( S^1 = \{ u \in \mathbb{C} : |u| = 1 \} \). Let \( \exp : \mathbb{R} \to S^1 \) denote the map \( \exp(\theta) = e^{i\theta} \), which is a smooth map. If \( a < b \in \mathbb{R} \) and \( b - a < 2\pi \), then \( \exp \) is a diffeomorphism from \((a, b)\) onto its image \( \exp(a, b) \subset S^1 \). This gives us charts \((U, \theta) = (\exp(a, b), (\exp|_V)^{-1})\). In the associated chart coordinates \((\tilde{U}, \tilde{\theta})\) on \( T^* S^1 \), any vector field has the form \( \tilde{X}(\tilde{\theta}) = f(\theta) \frac{d}{d\theta} \). Let us take \( f \equiv 1 \), so we have just the coordinate vector field \( \frac{d}{d\theta} \) on \( U \).

Now, let \((c, d)\) be any other open subset of \( \mathbb{R} \), and let \((V, \phi) = (\exp(c, d), (\exp|_V)^{-1})\). If \( U \cap V \neq \emptyset \), then for any point \( u \in U \cap V \) we have \( \theta = \theta(u) \) and \( \phi = \phi(u) \) defined so that \( e^{i\theta} = \exp(\theta) = \exp(\phi) = e^{i\phi} \). It follows that \( \phi = \theta + 2n\pi \) for some integer \( n \) (and since both \( \theta \) and \( \phi \) are smooth functions, \( n \) is constant on \( U \cap V \)). Hence \( \frac{d}{d\theta}|_u = \frac{d}{d\phi}|_u \). This means we can define a global vector field this way: it is defined to be \( \frac{d}{d\theta}|_u \) in any coordinate patch \((U, \theta)\) for \( U \) an open strict subset of \( S^1 \) where \( \theta \) is a right-inverse to \( \exp \). It is clearly smooth in any such patch, and the above argument shows it is well defined. Even though there is no global coordinate chart to define it, we still commonly denote it \( \frac{d}{d\theta} \).

Since a vector field is a smooth map \( M \to TM \), all our tools for smooth maps readily apply. Additionally, since we can (almost) think of a vector field as a function (taking values in a vector space), some of those relevant tools also apply, such as Proposition 2.29.

**Proposition 4.5.** Let \( A \subset M \) be a closed set, and let \( X \) be a smooth vector field on \( A \) (meaning that the map \( X : A \to TM \) is smooth in the sense of Definition 2.28). For any \( p \in A \), there is an open neighborhood \( V \) and a smooth extension of \( X|_{V \cap A} \) to \( U \). If \( V \) is any open neighborhood of \( p \) in \( M \), there is a smooth vector field \( \tilde{X} \in \mathcal{X}(M) \) such that \( \tilde{X}|_A = X \) and \( \text{supp}\tilde{X} \subset U \).

**Proof.** This is an exercise on Homework 3. \( \square \)

**Corollary 4.6.** Let \( M \) be a smooth manifold and \( p \in M \). Let \( v \in T_p M \). There is a smooth vector field \( X \in \mathcal{X}(M) \) such that \( X|_p = v \).

**Proof.** The set \( \{p\} \) is closed. Let \((U, \varphi)\) be any chart at \( p \); then \( v = \sum_{j=1}^n v^j \frac{\partial}{\partial x^j}|_p \) for some coefficients \( v^j \). The constant-coefficient vector field (given by the same sum) is smooth on \( U \), and is equal to \( v \) at \( \{p\} \). Hence, by Proposition 4.5, there is a global vector field \( X \in \mathcal{X}(M) \) which agrees with \( X \) on \( \{p\} \), i.e. \( X(p) = v \) as required. \( \square \)

Proposition 4.5 works here because we can use partitions of unity and “add and multiply in the codomain”. That is, even though \( TM \) is not typically a vector space, it is still possible to use vector space operations there locally. Another consequence of this is that \( \mathcal{X}(M) \), the space of smooth sections of \( TM \), is a vector space – in fact, it is a \( C^\infty(M) \)-module.

**Proposition 4.7.** Let \( M \) be a smooth manifold. Given \( X, Y \in \mathcal{X}(M) \), define \( X + Y \) to be the vector field \( (X + Y)|_p = X(p) + Y(p) \); for \( f \in C^\infty(M) \), define \( fX \) to be the vector field \( (fX)|_p = f(p)X_p \). These are smooth vector fields, and the operations so-defined make \( \mathcal{X}(M) \) into a \( C^\infty(M) \)-module.
Proof. $X + Y$ and $fX$ are defined to be vector fields (for each $p$ the result is a vector field at $p$ because $T_pM$ is a vector space, and they are defined to be based at $p$). Writing in local coordinates shows that they are smooth vector fields (because $X$, $Y$, and $f$ are smooth). Verifying the module properties is a trivial exercise. \[ \square \]

So, for example, (4.1) could be thought of not only as a decomposition at each point, but as an equality between vector fields in $\mathcal{X}(U)$:

$$X = \sum_{j=1}^{n} X_j^j \frac{\partial}{\partial x^j}$$

where $X_j^j$ is the smooth function $p \mapsto X_p^j$, and $\frac{\partial}{\partial x^j}$ is the smooth vector field $p \mapsto \frac{\partial}{\partial x^j}|_p$.

4.2. Frames. Continuing the theme of vector space terminology applied to the tangent bundle, we have the following.

**Definition 4.8.** Let $M$ be a smooth manifold. A collection $X_1, \ldots, X_k \in \mathcal{X}(M)$ of vector fields is called **linearly independent** if $\{X_1(p), \ldots, X_k(p)\}$ is a linearly independent set of vectors in $T_pM$ for each $p \in M$. (This is, of course, only possible if $k \leq \dim M$.) If $\text{span}\{X_1(p), \ldots, X_k(p)\} = T_pM$ for each $p$, we say that $X_1, \ldots, X_k$ span the tangent bundle. A frame for $M$ is a spanning set of linearly independent vector fields.

**Example 4.9.** (1) The vector fields $\{\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\}$ form a frame for $\mathbb{R}^n$. More generally, if $(U, \varphi)$ is any chart on an $n$-manifold, the coordinate vector fields (with the same notation as above) form a frame for $U$.

(2) The vector field $\frac{d}{du}$ on $\mathbb{S}^1$ is a frame. Here is a nice way to view it (acting on $\mathbb{S}^1 \subset \mathbb{C}$) that will be useful in the next example. Writing the vector field in the original framework (i.e. applying $\varphi^{-1}$ at each point $u \in \mathbb{S}^1$), this vector field is simply $X_u = iu$ for $u \in \mathbb{S}^1$.

(3) The 3-sphere $\mathbb{S}^3$ is defined to be the set of unit-norm points in $\mathbb{R}^4$. There is a nice algebra structure on $\mathbb{R}^4$ called the **quaternions**: we realize $\mathbb{R}^4$ as $\text{span}_\mathbb{R}\{1, i, j, k\}$ where $i, j, k$ are linearly independent and satisfy $ij = -ji = k$, $jk = -kj = i$, and $ki = -ik = j$. Extending this product by linearity defines an associative, non-commutative division algebra, the quaternions $\mathbb{H}$ (which is an example of a Clifford algebra).

We define three vector fields $X, Y, Z$ on $\mathbb{R}^4 \cong \mathbb{H}$ by $X(x) = ix$, $Y(x) = jx$, and $Z(x) = kx$ (where we suppress the “based at $x$” pair notation). These can be written explicitly in terms of the standard $\mathbb{R}^4$ coordinates $x = (x^1, x^2, x^3, x^4)$ as

$$X(x) = \begin{bmatrix} -x^2 \\ x^1 \\ -x^4 \\ x^3 \end{bmatrix}, \quad Y(x) = \begin{bmatrix} -x^3 \\ x^4 \\ x^1 \\ -x^2 \end{bmatrix}, \quad Z(x) = \begin{bmatrix} -x^4 \\ -x^3 \\ x^2 \\ x^1 \end{bmatrix}.$$  

These are linear functions of the coordinates, and so are smooth. One can readily check that the four vectors $\{u, X(u), Y(u), Z(u)\}$ form an orthonormal basis for $\mathbb{R}^4$ whenever $|u| = 1$. This shows that the vectors $X(u), Y(u), Z(u)$ are tangent to $\mathbb{S}^3$ at $u$, and are orthonormal, so linearly independent. Since $\dim T_u\mathbb{S}^3 = 3$, it follows that $X(u), Y(u), Z(u)$ span $T_u\mathbb{S}^3$. Thus, these three vector fields $X, Y, Z$ form a frame for $\mathbb{S}^3$. 


As part (1) above points out, it is always possible to find a local frame – i.e. a frame for an open neighborhood of any given point. But global frames are harder to come by. One might hope that something like the trick in part (3) of the example should work on any sphere, but this is decidedly false. While something like this works on $\mathbb{S}^7$, it doesn’t work on any sphere $\mathbb{S}^n$ unless $n \in \{1, 3, 7\}$. For example, the Hairy Ball Theorem asserts that, if $n$ is even, there does not exist a smooth (or even continuous) vector field on $\mathbb{S}^n$ that vanishes nowhere. (For an elementary proof of this fact, see the beautiful paper [2].)

A smooth manifold that admits a global (smooth) frame is called parallelizable. Parallelizable manifolds have trivial tangent bundles.

**Proposition 4.10.** Let $M^n$ be parallelizable. Then there is a diffeomorphism $F: M \times \mathbb{R}^n \to TM$ such that $F(p, v) \in T_pM$ for every $p \in M$. In particular, $TM$ is trivial.

**Proof.** Fix a frame $\{X_1, \ldots, X_n\}$ for $M$, and define $F$ by

$$F(p, (v^1, \ldots, v^n)) = v^1X_1(p) + \cdots + v^nX_n(p).$$

Then $F(p, v) \in T_pM$ as claimed. Note that $F$ is clearly a bijection: as $\{X_1, \ldots, X_n\}$ is a frame, every vector $X_p \in T_pM$ can be written uniquely in the form $v^1X_1(p) + \cdots + v^nX_n(p)$ for some $v = (v^1, \ldots, v^n) \in \mathbb{R}^n$, and so $F^{-1}(X_p) = (p, v)$. We need only show that $F$ and $F^{-1}$ are smooth. This is an exercise in writing vector fields in local coordinates (and changing coordinates), and is left to the reader. \(\square\)

The converse of **Proposition 4.10** is also true in the following specific sense. We call $TM$ a trivial vector bundle if there is a diffeomorphism $F: M \times \mathbb{R}^n \to TM$ with the properties

1. $F(p, v) \in T_pM$ for all $p \in M$ and $v \in \mathbb{R}^n$, and
2. for each $p \in M$, $F|_{T_pM}$ is a vector space isomorphism $\mathbb{R}^n \to T_pM$.

Such a diffeomorphism is called a trivialization of $TM$. **Proposition 4.10** really shows that if $M$ is parallelizable then $TM$ is a trivial vector bundle, and the converse is also true: if $F$ is a trivialization, then selecting any basis $\{e^1, \ldots, e^n\}$ of $\mathbb{R}^n$, the vector fields $X_j(p) = F(p, e_j)$ are a frame for $M$.

So, in particular, the Hairy Ball Theorem shows that $T\mathbb{S}^n$ is not a trivial vector bundle whenever $n$ is even, while Example 4.9 shows that $T\mathbb{S}^1$ and $T\mathbb{S}^3$ are trivial vector bundles. The more delicate question of whether it might still be true that $T\mathbb{S}^n \cong \mathbb{S}^n \times \mathbb{R}^n$ (via a diffeomorphism that does not preserve the bundle structure) is more delicate. The answer is still no for $n \not\in \{1, 3, 7\}$, but we don’t have the technology to prove it just now.

### 4.3. Derivations

A vector $X_p \in T_pM$ is a derivation at $p$. In particular, it is a certain kind of function $X_p: C^\infty(M) \to \mathbb{R}$. Denote this function space as $X_p \in \text{Fun}(C^\infty(M), \mathbb{R})$. Thus, a (rough) vector field is a function $X$ that takes values in $\text{Fun}(C^\infty(M), \mathbb{R})$:

$$X \in \text{Fun} \left( M, \text{Fun}(C^\infty(M), \mathbb{R}) \right).$$

For any sets $A, B, C$, there is a natural identification

$$\zeta: \text{Fun} \left( A, \text{Fun}(B, C) \right) \to \text{Fun} \left( B, \text{Fun}(A, C) \right)$$

given by

$$(\zeta(F)(b))(a) = (F(a))(b).$$

(It is immediate to check that this is a bijection; it will also preserve any reasonable extra structure on the spaces $A, B, C$, as we will see.) So, in our case, let $X$ be a (rough) vector field on $M$. Then
\( \zeta(X) \in \text{Fun}(C^\infty(M), \text{Fun}(M, \mathbb{R})) \): \( X \) can be viewed as a function which eats smooth functions \( f \in C^\infty(M) \) and spits out real-valued functions on \( \mathbb{R} \). The action is

\[
(\zeta(X)(f))(p) = X_p(f).
\]

A good notation for functions that allows us to omit the explicit reference to the variable \( p \) is to just represent it as an empty set of parentheses (\()) when needed; so we have \( \zeta(X)(f) = X(\text{\(f\)}). \)

The defining properties of \( X_p \) are that it is linear and obeys the product rule at the point \( p \). It then follows that, for \( \alpha, \beta \in \mathbb{R} \) and \( f, g \in C^\infty(M) \),

\[
\zeta(X)(\alpha f + \beta g) = X(\alpha f + \beta g) = \alpha X(\text{\(f\)}) + \beta X(\text{\(g\)}).
\]

That is to say: both the domain \( C^\infty(M) \) and codomain \( \text{Fun}(M, \mathbb{R}) \) of \( \zeta(X) \) are linear spaces, and \( \zeta(X) \) is a linear map. As for the derivation property, we have

\[
\zeta(X)(f g)(p) = X_p(f g) = f(p)X_p(g) + g(p)X_p(f) = f(p)(\zeta(X)(g))(p) + g(p)(\zeta(X)(f))(p).
\]

The expression on the right is a sum of products of functions \( M \to \mathbb{R} \); so we have the identity

\[
\zeta(X)(f g) = f\zeta(X)(g) + g\zeta(X)(f).
\]

**Definition 4.11.** A rough derivation on \( M \) is a linear function \( D : C^\infty(M) \to \text{Fun}(M, \mathbb{R}) \) with the property that \( D(f g) = fD(g) + gD(f) \) for all \( f, g \in C^\infty(M) \). If \( D \) takes values in \( C^\infty(M) \), we call it a smooth derivation, or just a derivation.

Our discussion above thus shows that if \( X \) is any (possibly rough) vector field on \( M \), then \( \zeta(X) \) is a rough derivation. From now on, we will abuse notation and drop the \( \zeta \), and identify the vector field \( X \) as this rough derivation: the action of the vector field is \( X(f)(p) = X_p(f) \) (where we now stop writing the second set of parentheses since we are identifying \( X(f)(p) = X(p(f)) \).

Now, we know that each tangent vector \( X_p \) is locally-defined: the value \( X_p(f) \) depends only on the germ of \( f \) at \( p \). This gives us a restriction operation on vector fields that does not make sense in general for functions defined on \( C^\infty(M) \). If \( U \subseteq M \) is an open subset, then we have a vector field \( X|_U \) which is a (rough) derivation \( C^\infty(U) \to \text{Fun}(U, \mathbb{R}) \). Its action (cf. Proposition 3.11) is

\[
X|_U(g)(p) = X_p(\tilde{g})
\]

where \( \tilde{g} \) is any smooth function on \( M \) whose restriction to some neighborhood of \( p \) agrees with \( g \) there. (Here we witness the identification of \( T_pU \) with \( T_pM \).) What this means is that, viewing \( X \) as a rough derivation, its domain is not really \( C^\infty(M) \); it is the “sheaf of germs of \( f \)” (an object which formally encodes the notion of making two functions equivalent if they agree on a neighborhood of each point, in a way that doesn’t mean the two functions are equal). In particular, it follows that, for any \( f \in C^\infty(M) \) and any open set \( U \subseteq M \),

\[
X|_U(f|_U) = X(f|_U).
\]

With this in hand, it becomes easy to characterize when a (possibly rough) vector field is actually smooth.

**Proposition 4.12.** Let \( X : M \to TM \) be a possibly rough vector field, and identify it as a rough derivation as above. The following are equivalent.

(a) \( X \) is smooth.
(b) \( X(f) \in C^\infty(M) \) for every \( f \in C^\infty(M) \).
(c) For every open set \( U \subseteq M \) and every \( g \in C^\infty(U) \), the function \( X|_U(g) \) is in \( C^\infty(U) \).

**Proof.** We prove the chain of implications \( (a) \implies (b) \implies (c) \implies (a) \).
• (a) \(\implies\) (b): If \(X\) is a smooth vector field, then for any point \(p \in M\), choose a chart \((U, \varphi)\) and write \(X_p = \sum_{j=1}^{n} X^j(p) \frac{\partial}{\partial x^j}|_p\). By Lemma 4.1, the coefficient functions \(X^j\) are smooth functions on \(U\). Thus, interpreting \(X\) as a rough derivation, we have for \(f \in C^\infty(M)\)

\[
X(f)(p) = X_p(f) = \sum_{j=1}^{n} X^j(p) \frac{\partial}{\partial x^j}|_p (f) = \sum_{j=1}^{n} X^j(p) \frac{\partial(f \circ \varphi^{-1})}{\partial x^j}(\varphi(p))
\]

which is a smooth function of \(p \in U\). Thus \(X(f)\) is smooth in a neighborhood of \(p\), for any \(p \in M\), proving that \(X(f) \in C^\infty(M)\).

• (b) \(\implies\) (c): We assume \(X\) is a smooth derivation, so in particular \(X(f) \in C^\infty(M)\) for all \(f \in C^\infty(M)\). Now, let \(U \subseteq M\) be open, and take any \(g \in C^\infty(U)\). Fix any \(p \in U\), and let \(\tilde{g} \in C^\infty(M)\) be a function that agrees with \(g\) on some neighborhood \(p\) (such a smooth extension can be constructed with a bump function, cf. the proof of Proposition 3.11). Then by (4.2), we have

\[
X|_U (g)(p) = X_p(\tilde{g}) = X(\tilde{g})(p)
\]

and since \(X(\tilde{g})\) is smooth by assumption, this shows \(X|_U (g)\) is smooth at \(p\). This holds for every \(p \in U\), so \(X|_U (g) \in C^\infty(U)\).

• (c) \(\implies\) (a) Let \(p \in M\), and let \((U, \varphi)\) be a chart at \(p\). By assumption, for any \(g \in C^\infty(U)\), \(X|_U (g)\) is smooth at \(p\). In particular, denoting \(\varphi(p) = (x^1(p), \ldots, x^n(p))\), the functions \(x^j\) are \(C^\infty\) on \(U\), and so \(X|_U (x^j)\) are smooth at \(p\). But we can compute these functions from the coordinate representation of the vector field \(X\):

\[
X|_U (x^j)(p) = X_p(x^j) = \sum_{k=1}^{n} X^k(p) \frac{\partial}{\partial x^k}|_p (x^j) = X^j(p).
\]

Hence, the conclusion is that the component functions \(X^j\) are smooth at \(p\) for each \(p \in U\). By Lemma 4.1, it follows that \(X\) is a smooth vector field.

\[\square\]

Thus, if \(X \in \mathcal{X}(M)\) is a smooth vector field, viewed as a rough derivation \(X: C^\infty(M) \to \text{Fun}(M, \mathbb{R})\), it is actually a derivation \(C^\infty(M) \to C^\infty(M)\). This turns out to be an equivalence. For the precise statement, it is useful to return to the \(\varsigma\) notation one more time.

**Theorem 4.13.** Let \(M\) be a smooth manifold. A function \(D: C^\infty(M) \to C^\infty(M)\) is a derivation if and only if \(D = \varsigma(X)\) for some \(X \in \mathcal{X}(M)\).

**Proof.** Proposition 4.12 shows that if \(X \in \mathcal{X}(M)\) then \(\varsigma(X)\) is a smooth derivation, leaving us only to prove the converse. Thus, given a smooth derivation \(D: C^\infty(M) \to C^\infty(M)\), define a vector field \(X: p \mapsto X_p\) the only way possible: \(X_p(f) = (D(f))(p)\) (which is precisely to say that \(\varsigma(X) = D\)). It is immediate that, since \(D\) is a derivation, \(f \mapsto (D(f))(p)\) is a derivation at \(p\), so \(X_p \in T_p M\), defining a rough vector field. Furthermore, since \(X_p(f) = (D(f))(p)\) is (by assumption) a smooth function of \(p\), it follows from Proposition 4.12 that \(X\) is a smooth vector field, concluding the proof.

\[\square\]

Henceforth, we will identify \(\mathcal{X}(M)\) with the space of smooth derivations on \(M\) (sometimes denoted \(\text{Der}(M)\)).
4.4. **Push-Forwards.** Let $M, N$ be smooth manifolds, with a smooth map $F : M \to N$. If $p \in M$ and $v \in T_p M$, then we transport $v$ to a vector $w = dF_p(v) \in T_{F(p)} N$. We might then expect that we can do the same to push forward a vector field $X \in \mathcal{X}(M)$ to a vector field on $N$. The naïve approach doesn’t quite work, however. The vector $dF_p(X(p))$ it tangent to $F(p) \in N$, so the only possible choice would be to define the pushed-forward vector field as $Y(q) = dF_p(X(p))$ where $q = F(p)$. But this relation may still hold between two vector fields. When it does, we say $X$ and $Y$ are $F$-related. Note that $F$ is the unique smooth vector field on $N$. But this relation does not generally define a vector field $Y$ from $X$. But this relation may still hold between two vector fields. When it does, we say $X$ and $Y$ are $F$-related. Since we now know we can view a smooth vector field as a derivation, let us translate the notion of $F$-relation in that language.

**Proposition 4.14.** Let $F : M \to N$ be a smooth map, $X \in \mathcal{X}(M)$, and $Y \in \mathcal{X}(N)$. Then $X$ and $Y$ are $F$-related iff

$$X(f \circ F) = Y(f) \circ F, \quad \forall f \in C^\infty(N).$$

**Proof.** Simply compute that, for any $p \in M$,

$$X(f \circ F)(p) = X_p(f \circ F) = dF_p(X_p)(f),$$

while

$$Y(f) \circ F(p) = Y(f)(F(p)) = Y_{F(p)}(f).$$

\[\square\]

**Example 4.15.** Let $X$ be the (global) coordinate vector field on $\mathbb{R}$ (day $X = \frac{d}{dt}$ with global coordinate $t$ on $\mathbb{R}$). Let $F : \mathbb{R} \to \mathbb{R}^2$ be the smooth map $F(t) = (\cos t, \sin t)$. Note that the image of $F$ is just the unit circle $S^1$, which amply shows how there can be no (unique) way to transport $X$ to all of $\mathbb{R}^2$ via $F$. However, there are plenty of $F$-related vector fields $Y$ to $X$. The most canonical one is

$$Y = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x};$$

the curl field. (Note that $Y|_{S^1}$ is the vector field $\frac{d}{dt}$ of Example 4.4.) To see this, we just compute: for any $f \in C^\infty(\mathbb{R}^2)$,

$$X(f \circ F)(t) = \frac{d}{dt}(f(\cos t, \sin t)) = \frac{\partial f}{\partial x}(\cos t, \sin t) \cdot (-\sin t) + \frac{\partial f}{\partial y}(\cos t, \sin t) \cdot (\cos t) = (Y f)(F(t)).$$

So $X$ and $Y$ are $F$-related. Note, however, that since $F$ can only “see” what happens to $Y$ on $S^1$, if we take any smooth vector $Y'$ field on $\mathbb{R}^2$ that agrees with $Y$ on $S^1$, then $X$ and $Y'$ will also be $F$-related.

The one case where there is a unique $F$-related vector field to a given $X \in \mathcal{X}(M)$ is when $F$ is a diffeomorphism.

**Proposition 4.16.** Let $M, N$ be smooth manifolds and let $F : M \to N$ be a diffeomorphism. Let $X \in \mathcal{X}(M)$. Define $F_*(X) \in \mathcal{X}(N)$ by

$$F_*(X)(q) = dF_{F^{-1}(q)}(X_{F^{-1}(q)}). \quad (4.3)$$

Then $F_*(X)$ is the unique smooth vector field on $N$ that is $F$-related to $X$. 

The vector field $F_*(X)$ is called the **push-forward** of $X$ by $F$. One might think of $Y$ as “a push forward of $X$ by $F$” whenever $F$ is a smooth map and $X, Y$ are $F$-related, but this is not uniquely defined unless $F$ is a diffeomorphism.

**Proof.** Let $Y_q = F_*(X)(q)$ be the (a priori rough) vector field defined by (4.3), and let $p = F^{-1}(q)$. Then (4.3) the statement $Y_q = dF_p(X_p)$, which is precisely the statement that $X$ and $Y$ are $F$-related. This shows that $Y$ is the unique (rough) vector field that is $F$-related to $X$. It remains only to show that $Y$ is in fact smooth. But by definition it is the composition

$$N \xrightarrow{F^{-1}} M \xrightarrow{X} TM \xrightarrow{dF} TN$$

which is a composition of three smooth maps, and is hence smooth. □

**Corollary 4.17.** Let $F: M \to N$ be a diffeomorphism, and let $X \in \mathcal{X}(M)$. Then for $f \in C^\infty(N)$,

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

Push forwards, and more generally $F$-relatedness, will be how we measure “naturality” of transformations of vector fields. Proposition 4.16 shows the unique way to push a vector field along a diffeomorphism, which shows that $\mathcal{X}(M)$ and $\mathcal{X}(N)$ are naturally isomorphic whenever $M$ and $N$ are diffeomorphic. More generally, if we have a smooth map $F: M \to N$ between manifolds, we will generally understand the behavior of a vector field on $M$ to be “the same” as a vector field on $N$ if they are $F$-related.

4.5. **Lie Brackets.** We now think of vector fields in $\mathcal{X}(M)$ as derivations: linear operators on $C^\infty(M)$ that satisfy the product rule. Locally, we know such operators can be expressed as first-order differential operators. Thus, the composition of two vector fields will, locally, be a second-order differential operator. Such operators are not derivations. Indeed, let us compute in general for two derivations $X, Y \in \text{Der}(M)$: for $f \in C^\infty(M)$,

$$XY(fg) = X(f \cdot Y(g) + g \cdot Y(f)) = [f \cdot XY(g) + Y(g)X(f)] + [g \cdot XY(f) + Y(f)X(g)]$$

$$= f \cdot XY(g) + g \cdot XY(f) + X(f)Y(g) + X(g)Y(f).$$

So, we see that the defect of $XY$ from being a derivation is

$$XY(fg) - [f \cdot XY(g) + g \cdot XY(f)] = X(f)Y(g) + X(g)Y(f).$$

This is typically not 0; it happens only when (viewed as maps $M \to TM$) the supports supp $X$ and supp $Y$ are disjoint.

So we cannot usually compose derivations to get a derivation. But we can take their **Lie bracket:**

$$[X, Y] \equiv XY - YX. \quad (4.4)$$

**Lemma 4.18.** Let $X, Y \in \mathcal{X}(M)$. Then $[X, Y] \in \mathcal{X}(M)$.

**Proof.** For any linear operator $A: C^\infty(M) \to C^\infty(M)$, let $\Gamma_A: C^\infty(M) \times C^\infty(M) \to C^\infty(M)$ measure the defect of $A$ from being a derivation,

$$\Gamma_A(f, g) = A(fg) - [f \cdot A(g) + g \cdot A(f)].$$

Then $A \in \mathcal{X}(M)$ if and only if $\mathcal{D}_A \equiv 0$. As computed above, we have

$$\Gamma_{XY}(f, g) = X(f)Y(g) + X(g)Y(f). \quad (4.5)$$
Now, \( A \mapsto \Gamma_A(f, g) \) is linear, and so we have
\[
\Gamma_{[X,Y]} = \mathcal{D}_{XY} - \mathcal{D}_{YX}.
\]
Subbing in (4.5) yields, for all \( f, g \in C^\infty(M) \),
\[
\Gamma_{[X,Y]}(f, g) = [X(f)Y(g) + X(g)Y(f)] - [Y(f)X(g) + Y(g)X(f)] = 0.
\]
\[\square\]

**Example 4.19.** Consider the following two vector fields on \( \mathbb{R}^2 \):
\[
X = \frac{\partial}{\partial x}, \quad Y = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}.
\]
We will call these two vector fields \( X \) = “drive east” and \( Y \) = “steer”. Let’s compute their bracket: for \( f \in C^\infty(\mathbb{R}^2) \),
\[
[X, Y](f) = \frac{\partial}{\partial x} \left( x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) f - \left( x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \frac{\partial}{\partial x} f
\]
\[
= \frac{\partial}{\partial x} (xf_y - yf_x) - (xf_y - yf_x)
\]
\[
= fy + xfx_y - yfx_x - (xf_y - yf_x) = \frac{\partial}{\partial y} f.
\]
That is: the bracket of “drive east” and “steer” is “drive north”. This is why parallel parking works! More in this later.

The Lie bracket of vector fields is a genuinely new operation: it is not just a new coordinate-free version of some vector operation from vector calculus. Let us compute it in local coordinates.

**Proposition 4.20.** Let \( M \) be a smooth manifold, let \( X, Y \in \mathcal{X}(M) \) be vector fields, and let \( (U, \varphi) \) be a coordinate chart with \( \varphi = (x^1, \ldots, x^n) \). Express the vector fields in local coordinates (as usual)
\[
X|_U = \sum_{j=1}^n X^j \frac{\partial}{\partial x^j}, \quad Y|_U = \sum_{j=1}^n Y^j \frac{\partial}{\partial x^j}.
\]
Then, in local coordinates,
\[
[X, Y]|_U = \sum_{j,k=1}^n \left( X^j \frac{\partial Y^k}{\partial x^j} - Y^j \frac{\partial X^k}{\partial x^j} \right) \frac{\partial}{\partial x^k}.
\]

**Proof.** This is an elementary (if tedious) computation. Fix \( f \in C^\infty(M) \). To save space, denote \( \frac{\partial}{\partial x^k} f = f_k \), and \( \frac{\partial}{\partial x^j} \frac{\partial}{\partial x^k} f = f_{kj} \). Since \( f \) is smooth, \( f_{kj} = f_{jk} \) as usual. Then (being clever about naming indices) we have
\[
[X, Y]|_U f = \left( \sum_{j=1}^n X^j \frac{\partial}{\partial x^j} \right) \left( \sum_{k=1}^n Y^k \frac{\partial}{\partial x^k} \right) f - \left( \sum_{j=1}^n Y^j \frac{\partial}{\partial x^j} \right) \left( \sum_{k=1}^n X^k \frac{\partial}{\partial x^k} \right) f
\]
\[
= \sum_{j,k=1}^n \left( X^j \frac{\partial}{\partial x^j} (Y^k f_k) - Y^j \frac{\partial}{\partial x^j} (X^k f_k) \right).
\]
The term inside the brackets expands (using the product rule) to
\[
X^j \frac{\partial}{\partial x^j} (Y^k f_k) - Y^j \frac{\partial}{\partial x^j} (X^k f_k) = X^j \frac{\partial Y^k}{\partial x^j} f_k + X^j Y^k f_{kj} - Y^j \frac{\partial X^k}{\partial x^j} f_k - Y^j X^k f_{kj}.
\]
Grouping the first- and second-order terms into two sums, this gives
\[
[X,Y]|_U f = \sum_{j,k=1}^{n} \left( X_j \frac{\partial Y_k}{\partial x^j} f_k - Y_j \frac{\partial X_k}{\partial x^j} f_k \right) + \sum_{j,k=1}^{n} (X_j Y_k f_{kj} - Y_j X_k f_{kj}).
\]

Finally, break up the second sum and reverse the dummy indices:
\[
\sum_{j,k=1}^{n} X_j Y_k f_{kj} - \sum_{j,k=1}^{n} Y_j X_k f_{kj} = \sum_{j,k=1}^{n} X_j Y_k f_{kj} - \sum_{j,k=1}^{n} Y_k X_j f_{jk} = 0,
\]
where the final equality comes from the identity \( f_{jk} = f_{kj} \).

In particular, if all the coefficients \( X_j \) and \( Y_j \) are constant, then \( [X,Y]|_U = 0 \). (A special case of this is when \( X = \frac{\partial}{\partial x^j} \) and \( Y = \frac{\partial}{\partial x^k} \) for some fixed \( j, k \); then the fact that \( [X,Y] = 0 \) is just the statement that \( f_{jk} = f_{kj} \) for smooth \( f \), which we used in the proof.) Whenever it happens that \( [X,Y] = 0 \), we say that the vector fields \( X \) and \( Y \) commute. Thus, coordinate vector fields (and in general constant coefficient vector fields) commute.

What does it mean for vector fields to commute? More generally: what does \( [X,Y] \) really mean? Proposition 4.20 shows that it is a new vector field with components built out of the components of \( X \) and \( Y \) together with their derivatives. In a sense, the goal of the next chapter is to understand what this really means.

First, let us summarize some algebraic properties of the Lie bracket that will be very useful for computations.

**Proposition 4.21.** Let \( M \) be a smooth manifold, and let \( X, Y, Z \in \mathfrak{X}(M) \). The Lie bracket satisfies the following properties.

(a) **Bilinearity:** for \( a, b \in \mathbb{R} \),
\[
[aX + bY, Z] = a[X,Z] + b[Y,Z],
\]
\[
\]

(b) **Antisymmetry:**
\[
[X, Y] = -[Y, X].
\]

(c) **Jacobi Identity:**
\[
[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.
\]

(d) For \( f, g \in C^\infty(M) \),
\[
[f \cdot X, g \cdot Y] = fg \cdot [X,Y] + (f \cdot Xg)Y - (g \cdot Yf)X.
\]

**Proof.** Items (a) and (b) are immediate from the definition. Item (c) is a simple if tedious computation that is left to the bored reader. The mysterious item (d) is left as a homework exercise. □

Let us take a moment to examine the Jacobi identity. In one sense, it is just a mnemonic for how to “associate” three elements (the Lie bracket is non-associative). To try to make a little more sense of it, we introduce some (important) notation. Given a vector field \( X \in \mathfrak{X}(M) \), denote by \( \text{ad}_X : \mathfrak{X} \to \mathfrak{X} \) the linear operator
\[
\mathcal{L}_X(Y) = [X,Y].
\]

Then we can rewrite the Jacobi identity as follows: first, using antisymmetry twice, write it as \( [X, [Y, Z]] = [[X,Y], Z] + [Y, [X,Z]] \). This can be written as
\[
\mathcal{L}_X([Y,Z]) = [\mathcal{L}_X(Y), Z] + [Y, \mathcal{L}_X(Z)].
\]
That is: if we take $\mathcal{X}(M)$ to be a (non-associative) algebra (called a Lie algebra), where the product is $[\cdot,\cdot]$, then the Jacobi identity says precisely that $L_X$ is a derivation on $\mathcal{X}(M)$. This suggests that we can interpret $L_X$ as a kind of “derivative with respect to $X$”. We will make this precise in the next chapter. We will also return to Proposition 4.21(d) in the next chapter, where it will also make sense as a statement about the Lie bracket being a certain kind of derivative.

Let us conclude this section (and chapter) by noting that the Lie bracket is a “natural” object: it preserves $F$-relation for any smooth map $F$.

**Proposition 4.22.** Let $M,N$ be smooth manifolds and let $F: M \to N$ be a smooth map. If $X_1,X_2 \in \mathcal{X}(M)$, $Y_1,Y_2 \in \mathcal{X}(N)$, and if $(X_j,Y_j)$ are $F$-related for $j = 1,2$, then $[X_1,X_2]$ and $[Y_1,Y_2]$ are $F$-related.

**Proof.** We use the $F$-relation twice to compute, for any $f \in C^\infty(N)$,

$$X_1 X_2 (f \circ F) = X_1 ((Y_2(f) \circ F) = (Y_1 Y_2(f)) \circ F, \quad \text{and}$$

$$X_2 X_1 (f \circ F) = X_2 (Y_1(f) \circ F) = (Y_2 Y_1(f)) \circ F.$$

Thus

$$[X_1, X_2] (f \circ F) = X_1 X_2 (f \circ F) - X_2 X_1 (f \circ F) = (Y_1 Y_2(f)) \circ F - (Y_2 Y_1(f)) \circ F = ([Y_1, Y_2](f)) \circ F. \quad \square$$

**Corollary 4.23.** Let $M,N$ be diffeomorphic manifolds via diffeomorphism $F: M \to N$. Let $X_1,X_2 \in \mathcal{X}(M)$. Then $F_*[X_1,X_2] = [F_*X_1,F_*X_2]$. 

5. Flows

5.1. Integral Curves. Let $M$ be a smooth manifold, and let $X \in \mathcal{X}(M)$ be a smooth vector field. So $X(p)$ is a tangent vector at $p$; in particular, it is the tangent vector to lots of curves passing through that point. If $\alpha$ is a smooth curve with the property that, for every point $p$ it passes through, its tangent vector is $X(p)$, it is called an integral curve of $X$. That is: the condition is that for $\alpha: (t_-, t_+) \to M$,
\begin{equation}
\dot{\alpha}(t) = X(\alpha(t)), \quad t \in (t_-, t_+).
\end{equation}

Example 5.1. Let $M = \mathbb{R}^2$, and take $X = \frac{\partial}{\partial x}$. Then it is easy to check that any curve of the form $\alpha(t) = p + te^1$, for any $p \in \mathbb{R}^2$, is an integral curve: $\dot{\alpha}(t) = e^1 \sim \frac{\partial}{\partial x} = X(\alpha(t))$. It is also easy to see that $\alpha(t) = p + te^1$ is the unique integral curve passing through $p$ at $t = 0$. Indeed, this is a statement about an ODE: for any curve $\alpha(t) = (x(t), y(t))$,
\begin{equation}
\frac{\partial}{\partial x} \bigg|_{\alpha(t)} X(\alpha(t)) = \dot{\alpha}(t) = d\alpha_t \left( \frac{d}{dt} \bigg|_t \right) = \frac{dx}{dt} \frac{\partial}{\partial x} \bigg|_{\alpha(t)} + \frac{dy}{dt} \frac{\partial}{\partial y} \bigg|_{\alpha(t)}.
\end{equation}

This sets up the ODE
\[ \frac{dx}{dt} = 1, \quad \frac{dy}{dt} = 0, \quad (x(0), y(0)) = p \]
and it is elementary to verify that the unique solution is $\alpha(t) = (x(t), y(t)) = p + te^1$.

In particular, we see that, for this example, given any $p \in M$, there is a unique integral curve $\alpha^p$ of $X$ that passes through $p$ at time $t = 0$, $\alpha^p(0) = p$; moreover, for any $p, q \in \mathbb{R}^2$, the images of $\alpha^p$ and $\alpha^q$ are either identical (if $p^1 = q^1$) or disjoint (otherwise).

Example 5.2. Let $M = \mathbb{R}^2$, and take $Y = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}$. Any integral curve $\alpha(t) = (x(t), y(t))$ is determined by
\[ x(t) \frac{\partial}{\partial y} \bigg|_{\alpha(t)} - y(t) \frac{\partial}{\partial x} \bigg|_{\alpha(t)} = X(\alpha(t)) = \dot{\alpha}(t) = \frac{dx}{dt} \frac{\partial}{\partial x} \bigg|_{\alpha(t)} + \frac{dy}{dt} \frac{\partial}{\partial y} \bigg|_{\alpha(t)}.
\]

In other words, we have the ODE
\[ \frac{dx}{dt} = -y(t), \quad \frac{dy}{dt} = x(t). \]

Subject to the constraint $\alpha(0) = (x(0), y(0)) = (x_0, y_0)$, the unique solution is
\[ \alpha(t) = (x(t), y(t)) = (x_0 \cos t - y_0 \sin t, x_0 \sin t + y_0 \cos t). \]

When $(x_0, y_0) = (0, 0)$, this is the constant curve $\alpha(t) = 0$; otherwise, it is a counter-clockwise circle. Once again, we see that, for every point $p = (x_0, y_0)$, there is a unique integral curve $\alpha^p$ with $\alpha^p(0) = p$, and the images of any two integral curves are either identical or disjoint.

Example 5.3. Let $M = \mathbb{R}^2$, and let $X = x^2 \frac{\partial}{\partial x}$. Then any integral curve $\alpha(t) = (x(t), y(t))$ satisfies the ODE
\[ \dot{x}(t) = x(t)^2, \quad \dot{y}(t) = 0. \]

The standard procedure is to “separate variables” and solve for $x$ by $\frac{dx}{x^2} = dt$. This only works if $x \neq 0$, of course; if the initial condition is of the form $\alpha(0) = (0, y_0)$, then we can easily check that the constant curve $\alpha(t) = (0, y_0)$ is the unique solution. Otherwise, if $x_0 \neq 0$, the unique solution with $\alpha(0) = (x_0, y_0)$ is
\[ x(t) = \frac{1}{1/x_0 - t}, \quad y(t) = y_0. \]
If $x_0 > 0$, the domain of this curve is $t \in (−∞, 1/x_0)$. As $t \uparrow 1/x_0$, the curve accelerates off to $∞$ along the line $y = y_0$. As $t \downarrow −∞$, the curve approaches the point $(0, y_0)$. If $x_0 < 0$, the behavior is the same (in the opposite direction).

Integral curves always exist locally.

**Proposition 5.4.** Let $M$ be a smooth manifold and $X \in \mathcal{X}(M)$. For each point $p \in M$, there exists an integral curve $\alpha$ for $X$ defined on some time interval $(-\epsilon, \epsilon)$ such that $\alpha(0) = p$.

**Proof.** Let $(U, \varphi)$ be a chart at $p$, with coordinate functions $\varphi = (x^1, \ldots, x^n)$. Write $\alpha$ in local coordinates as $\alpha(t) = (\alpha^1(t), \ldots, \alpha^n(t))$. (Technically we should write $\dot{\alpha}(t) = \varphi \circ \alpha(t)$, but at this point we will start dropping that extra notation.) Also write the vector field in local coordinates

$$X = \sum_{j=1}^{n} X^j \frac{\partial}{\partial x^j}.$$ 

Then (5.1) becomes the ODE

$$\frac{d\alpha^j}{dt} = X^j(\alpha^1(t), \ldots, \alpha^n(t)), \quad 1 \leq j \leq n.$$ 

By Theorem 0.22 (the Picard–Lindelöf Theorem), there is a unique smooth solution $\alpha(t)$ with $\alpha(0) = p$, for some time interval $(-\epsilon, \epsilon)$, proving the proposition. □

**Remark 5.5.** The Picard-Lindelöf theorem shows that the integral curve is unique in the chart $U$. We will need more work to show that it is unique globally on $M$.

The next few lemmas investigate how integral curves respond to transformations of their vector field.

**Lemma 5.6** (Dilation). Let $M$ be a smooth manifold, and $X \in \mathcal{X}(M)$. Let $\alpha: (t_-, t_+) \to M$ be an integral curve for $X$. For any $a \in \mathbb{R}$, the curve $\delta_a(\alpha): \{t: at \in (t_-, t_+)\} \to M$ defined by $\delta_a(\alpha)(t) = \alpha(at)$ is an integral curve for the vector field $aX$.

**Proof.** This is just a calculation: for any $t_0 \in \mathbb{R}$ with $at_0 \in (t_-, t_+)$, and any $f \in C^\infty(M)$,

$$\delta_a(\alpha)'(t_0)(f) = \frac{d}{dt} \bigg|_{t_0} f \circ \delta_a(\alpha)(t) = \frac{d}{dt} \bigg|_{t_0} (f \circ \alpha)(at) = a(f \circ \alpha)'(at_0) = a\alpha'(at_0)f = aX(\delta_a(\alpha))f.$$ 

□

**Lemma 5.7** (Translation). Let $M$ be a smooth manifold, and $X \in \mathcal{X}(M)$. Let $\alpha: (t_0, t_1) \to M$ be an integral curve for $X$. For any $b \in \mathbb{R}$, the curve $\tau_b(\alpha): (t_0 - b, t_1 - b) \to M$ defined by $\tau_b(\alpha)(t) = \alpha(t + b)$ is also an integral curve of $X$.

**Proof.** On Homework 3. □

**Lemma 5.8** (“Integral curves are natural”). Let $M, N$ be smooth manifolds with a smooth map $F: M \to N$. Let $X \in \mathcal{X}(M)$ and $Y \in \mathcal{X}(N)$. Then $X$ is $F$-related to $Y$ iff $F$ takes integral curves of $X$ to integral curves of $Y$ (i.e. for every integral curve $\alpha: (t_-, t_+) \to M$ of $X$, $F \circ \alpha: (t_-, t_+) \to N$ is an integral curve of $Y$).

**Proof.** Suppose $X, Y$ are $F$-related, i.e. $dF_p(X(p)) = Y(F(p))$ for all $p \in M$. Let $\alpha$ be an integral curve of $X$; thus $\alpha'(t) = X(\alpha(t))$ for all $t$. So

$$(F \circ \alpha)'(t)dF_{\alpha(t)}(\alpha'(t)) = dF_{\alpha(t)}(X(\alpha(t))) = Y(F(\alpha(t))) = Y(F \circ \alpha(t)),$$
showing that $F \circ \alpha$ is an integral curve of $Y$.

Conversely, suppose $F$ maps integral curves of $X$ to integral curves of $Y$. For any $p \in M$, choose an integral curve $\alpha$ of $X$ with $\alpha(0) = p$. Then

$$dF_p(X(p)) = dF_{\alpha(0)}(X(\alpha(0))) = dF_{\alpha(0)}(\alpha'(0)) = (F \circ \alpha)'(0).$$

By assumption, $F \circ \alpha$ is an integral curve of $Y$, which means that $(F \circ \alpha)'(0) = Y(F \circ \alpha(0)) = Y(F(p))$. This shows that $X, Y$ are $F$-related. \qed

5.2. Flows. In Examples 5.1 and 5.2 we saw vector fields on a smooth manifold $M$ whose integral curves had the following property:

Given any $p \in M$, there is a unique integral curve $\alpha_p : \mathbb{R} \to M$ with $\alpha_p(0) = p$.

As Example 5.3 shows, this property does not always hold: in particular, the curve may not exist for all time. Casting this aside for the moment, suppose $X$ is a vector field with unique integral curves $\{\alpha_p : p \in M\}$ all existing for all time. Then we can define a map

$$\theta : \mathbb{R} \times M \to M, \quad \theta(t, p) = \alpha_p(t).$$

For fixed $t \in \mathbb{R}$, let $\theta_t : M \to M$ be the map $\theta_t(p) = \theta(t, p)$. Note that $\theta_0(p) = \theta(0, p) = \alpha_p(0) = p$ by definition; so $\theta_0 = \text{Id}_M$. Now, let $s, t \in \mathbb{R}$, and consider $\theta(t + s, p)$. This is the point $\alpha_p(t + s)$: the point reached by the unique integral curve of $X$ starting at $p$ after time $t + s$. On the other hand, the translation Lemma asserts that $t \mapsto \alpha_p(t + s)$ is also an integral curve of $X$; it starts at $q = \alpha_p(s)$. That is, we have the relationship

$$\alpha_p(t + s) = \alpha_q(t), \quad \text{where} \quad q = \alpha_p(s).$$

Translating this into $\theta$-language, this says that $\theta(t + s, p) = \theta(t, q) = \theta(t, \alpha_p(s)) = \theta(t, \theta(s, p))$; or, more succinctly

$$\theta_{t+s} = \theta_t \circ \theta_s. \quad (5.2)$$

We can derive a number of consequences immediately from this.

- For each $t \in \mathbb{R}$, $\theta_t : M \to M$ is a bijection. Indeed, $\theta_t \circ \theta_{-t} = \theta_{-t} \circ \theta_t = \theta_0 = \text{Id}_M$, so $(\theta_t)^{-1} = \theta_{-t}$.
- Given any two integral curves $\alpha_p, \alpha_q$, either $\alpha_p(\mathbb{R}) = \alpha_q(\mathbb{R})$, or $\alpha_p(\mathbb{R}) \cap \alpha_q(\mathbb{R}) = \emptyset$. After all, if there is some point $r \in \alpha_p(\mathbb{R}) \cap \alpha_q(\mathbb{R})$, then there are times $t_0, s_0 \in \mathbb{R}$ so that $r = \alpha_p(t_0) = \alpha_q(s_0)$. This means that $\theta_{t_0}(p) = \theta_{s_0}(q)$. Applying (5.2), we have, for any $t$, $\theta_{t_0}(\theta_{s_0}(q)) = \theta_{t+s_0-t_0}(q)$.

I.e. $\alpha_p(t) = \alpha_q(t + s_0 - t_0) = (\tau_{s_0-t_0} \alpha_q)(t)$, so $\alpha_p = \tau_{s_0-t_0} \alpha_q$. The two curves are translations of each other, and so (since both are defined for all time) they have the same image, as claimed.

Motivated by this discussion, we define a flow as follows.

**Definition 5.9.** A (smooth) global flow on $M$ is a smooth function $\theta : \mathbb{R} \times M \to M$ for which, writing $\theta_t(p) = \theta(t, p)$, we have $\theta_0 = \text{Id}_M$ and (5.2) holds.

From the above discussion, we see that each map $\theta_t$ for fixed $t \in \mathbb{R}$ is a diffeomorphism of $M$: it is a smooth map with inverse $\theta_{-t}$, which is (by definition) also smooth. So a smooth global flow is a 1-parameter group of diffeomorphisms.

Given a global flow, we get a collection of curves $\{\theta^p : p \in M\}$ defined by $\theta^p(t) = \theta_t(p)$ for all $t \in \mathbb{R}$. (So $\theta^p$ is what we formerly called $\alpha_p$.) Since $\theta$ is smooth (in both variables), each curve $\alpha_p$
is a smooth curve. By the second item in the discussion above Definition 5.9, the images of these curves are either identical or disjoint, and they (of course) cover \( M \).

Our next proposition shows that any smooth global flow comes from a vector field as per the discussion leading up to the definition.

**Proposition 5.10.** Let \( \theta \) be a smooth global flow on \( M \). For each \( p \in M \), let \( \theta^p(t) = \theta(t, p) \), and define a vector \( X_p \in T_p M \) by

\[
X_p = \dot{\theta}^p(0).
\]

Then the rough vector field \( X : \theta \mapsto X_p \) is smooth, and for each \( p \), \( \theta^p \) is an integral curve of \( X \) starting at \( p \).

We call the vector field \( X \) above the **infinitesimal generator** of the flow \( \theta \).

**Proof.** Let \( f \in C^\infty(M) \). Then we calculate

\[
X(p)(f) = X_p(f) = (\theta^p(0))(f) = d(\theta^p)_p \left( \frac{d}{dt} \right)_{0}(f) = \frac{d}{dt} \bigg|_0 f(\theta^p(t)) = \frac{\partial}{\partial t} \bigg|_{(0,p)} f(\theta(t,p)).
\]

The function \( f \circ \theta \) is smooth, and hence so are its partial derivatives. This shows that \( X(f) \) is smooth, and so by Proposition 4.12, \( X \) is a smooth vector field.

Now, we must show that \( \dot{\theta}^p \) is an integral curve of \( X \) starting at \( p \). We have \( \theta^p(0) = \theta_0(p) = p \) by definition. Fix \( t_0 \in \mathbb{R} \). Let \( q = \theta^p(t_0) \). Then we can compute, for any \( f \in C^\infty(M) \),

\[
X(q)f = (\theta^0(0))(f) = \left. \frac{d}{dt} \right|_0 f \circ \theta^q(t) = \left. \frac{d}{dt} \right|_0 f(\theta_t(q)).
\]

But \( \theta_t(q) = \theta_t(\theta^p(t_0)) = \theta_t \circ \theta_{t_0}(p) = \theta_{t+t_0}(p) = \theta^p(t+t_0) \), and so

\[
X(q)f = \left. \frac{d}{dt} \right|_0 f(\theta^p(t+t_0)) = (\dot{\theta}^p(t_0))(f).
\]

That is: we have shown that, as derivations,

\[
X(\theta^p(t_0)) = \dot{\theta}^p(t_0), \quad t_0 \in \mathbb{R}.
\]

This concludes the proof that \( \theta^p \) is an integral curve of \( X \). \( \square \)

**Example 5.11.** Let \( M = \mathbb{R}^n \), and define \( \theta(t,p) = e^tp \), the dilation (by exponential time). This is a smooth map of both variables, with \( \theta_0(p) = p \) and \( \theta_{s+t}(p) = e^{s+t}p = e^se^tp = \theta_s \circ \theta_t(p) \); so \( \theta \) is a flow. Its infinitesimal generator \( E \) is the vector field

\[
E(p)(f) = \left. \frac{\partial}{\partial t} \right|_{(0,p)} f(\theta(t,p)) = \left. \frac{\partial}{\partial t} \right|_{(0,p)} f(e^tp) = \sum_{j=1}^n p^j e^t \frac{\partial f}{\partial x^j}(e^tp) \bigg|_{t=0} = \sum_{j=1}^n p^j \left. \frac{\partial}{\partial x^j} \right|_{(0,p)} f.
\]

That is: \( E(p) = \sum_{j=1}^n p^j \left. \frac{\partial}{\partial x^j} \right|_{(0,p)} \) is the Euler field of Example 4.3. In other words, the Euler field is the infinitesimal generator of dilations.

As Example 5.3 shows, not every vector field is the infinitesimal generator of a global flow. When this happens, we call the vector field **complete**; more on complete vector fields later. It turns out that the only part that fails is the definition of the integral curves for all time. To deal with this, we weaken the definition of a flow as follows.
Definition 5.12. Let $M$ be a manifold. A flow domain $\mathcal{D} \subseteq \mathbb{R} \times M$ is an open set for which each section $\mathcal{D}^p = \{ t \in \mathbb{R} : (t, p) \in \mathcal{D} \}$ is an open interval in $\mathbb{R}$ containing 0. A (smooth) flow on $M$ is a smooth function $\theta : \mathcal{D} \to M$ defined on some flow domain, with the properties:

- $\theta(0, p) = p$ for all $p \in M$, and
- for any $p \in M$, if $s \in \mathcal{D}^p$ and $t \in \mathcal{D}^{\theta(s, p)}$ satisfy $s + t \in \mathcal{D}^p$, then
  \[ \theta(t, \theta(s, p)) = \theta(t + s, p). \]

Some authors call such an object a local flow (to distinguish it from a global flow). As usual, we let $\theta_t(p) = \theta^p(t) = \theta(t, p)$; so $\theta^p$ is a curve defined on the open interval $\mathcal{D}^p$. For fixed $t$, the set of $p$ for which $\theta_t(p)$ makes sense is denoted $M_t$:
\[ M_t = \{ p \in M : (t, p) \in \mathcal{D} \}. \]

Example 5.13. In Example 5.3, we found the integral curve $\alpha_p$ of $x^2 \partial / \partial x$ starting at $p = (x_0, y_0)$ was defined on the time interval $(-\infty, 1/x_0)$ of $x_0 > 0$, on the interval $(1/x_0, \infty)$ if $x < 0$, and on $\mathbb{R}$ if $x_0 = 0$. So we define this as our flow domain:
\[ \mathcal{D} = \{ (t, (x, y)) \in \mathbb{R} \times \mathbb{R}^2 : x > 0, t < 1/x \} \cup \{ (t, (x, y)) \in \mathbb{R} \times \mathbb{R}^2 : x < 0, t > 1/x \} \cup \mathbb{R} \times \{ 0 \} \times \mathbb{R} \]
\[ = \{ (t, (x, y)) : xt < 1, y \in \mathbb{R} \}. \]

This is an open set. The definition as a union of three pieces shows that $\mathcal{D}^{(x,y)}$ for the three regions $x > 0$, $x < 0$, and $x = 0$. Drawing the picture, we see that $(\mathbb{R}^2)_t = \{ (x, y) : x < 1/t, y \in \mathbb{R} \}$ if $t > 0$, $(\mathbb{R}^2)_t = \{ (x, y) : x > 1/t, y \in \mathbb{R} \}$ if $t < 0$, and $(\mathbb{R}^2)_0 = \mathbb{R}^2$.

Putting the integral curves of the vector field together, the flow on this domain is
\[ \theta(t, (x, y)) = \left( \frac{1}{1/x - t}, y \right) \text{ for } x \neq 0, \text{ and } \theta(t, (0, y)) = (0, y). \]

This is a smooth function on this domain. We can easily check that it satisfies the flow properties (but we don’t need to, since this flow arose from the unique integral curves of a vector field).

Note that a smooth (local) flow defines a smooth vector field exactly as a global one does: if $\theta : \mathcal{D} \to M$ is a smooth flow, then for each $p$ there is an open neighborhood (in $\mathbb{R} \times M$) of $(0, p)$ contained in $\mathcal{D}$. Thus, the proof of Proposition 5.10 carries through without alteration, and we find that any flow has an infinitesimal generator defined by (5.3).

Given a flow $\theta$ on a domain $\mathcal{D}$, we can always restrict it to a smaller domain $\mathcal{D}' \subset \mathcal{D}$ (provided $\mathcal{D}'$ is still a flow domain). For example, we could take the global flows of Examples 5.1 and 5.2 and restrict them to some flow domain smaller than $\mathbb{R}^2$. Then the infinitesimal generator would still be the same vector field in each case – it is determined only by the behavior of the flow in any small tubular neighborhood of the section $\{ t = 0 \}$ in the flow domain. This highlights that one cannot recover the flow domain from the infinitesimal generator, as we’ve defined things. But we can fix this by insisting the flow domain be as large as possible.

Definition 5.14. A maximal flow is a flow $\theta$ on a flow domain $\mathcal{D}$ with the property that the function $\theta$ has no smooth extension to any flow domain larger than $\mathcal{D}$. A maximal integral curve of a vector field is an integral curve which has no smooth extension to an integral curve on a strictly larger interval.

With these definitions, vector fields and flows come into one-to-one correspondence.
Theorem 5.15 (Fundamental Theorem on Flows). Let $M$ be a smooth manifold, and let $X \in \mathcal{X}(M)$. There is a unique maximal flow $\theta : \mathcal{D} \to M$ whose infinitesimal generator is $X$. Moreover, the flow satisfies the following properties.

(a) For each $p \in M$, the curve $\theta^p$ is the unique maximal integral curve of $X$ starting at $p$.
(b) If $s \in \mathcal{D}^p$, then $\mathcal{D}^{\theta(s,p)} = \mathcal{D}^p - s = \{ t - s : t, s \in \mathcal{D}^p \}$.
(c) For each $t \in \mathbb{R}$, the set $M_t$ is open in $M$, and $\theta_t : M_t \to M_{-t}$ is a diffeomorphism, with inverse $\theta_{-t}$.

In order to prove Theorem 5.15, we will first show that every vector field has a unique maximal integral curve starting at any given point.

Lemma 5.16. Let $M$ be a smooth manifold, and let $X \in \mathcal{X}(M)$. For each $p \in M$, there is a unique maximal integral curve $\alpha_p$ with $\alpha_p(0) = p$.

Proof. For each $p \in M$, by Proposition 5.4 there is an integral curve $\alpha : (-\epsilon_p, \epsilon_p) \to M$ that starts at $p$. Now, let $\alpha$ and $\tilde{\alpha}$ be two integral curves of $X$ defined on intervals $(t_-, t_+)$ and $(\tilde{t}_-, \tilde{t}_+)$ (not necessarily containing $0$). Suppose the two time intervals intersect: $(t_-, t_+) \cap (\tilde{t}_-, \tilde{t}_+) = (a, b)$, and suppose there exists some $t_0 \in (a, b)$ with $\alpha(t_0) = \tilde{\alpha}(t_0)$. We will show that, in fact, $\alpha(t) = \tilde{\alpha}(t)$ for all $t \in (a, b)$. To do this, define $\mathcal{T} = \{ t \in (a, b) : \alpha(t) = \tilde{\alpha}(t) \}$. Then $t_0 \in \mathcal{T}$, so the set is nonempty. Now, fix any metric $d_M$ that metrizes the topology of $M$, and note that $\mathcal{T} = \{ t \in (a, b) : d_M(\alpha(t), \tilde{\alpha}(t)) = 0 \}$. The function $t \mapsto d_M(\alpha(t), \tilde{\alpha}(t))$ is continuous, and so $\mathcal{T}$ is the preimage of $\{0\}$ under a continuous map; it is therefore a closed set.

Now, let $t_1$ is any point in $\mathcal{T}$. In a chart at $\alpha(t_1) = \tilde{\alpha}(t_1) \in M$, the condition that $\alpha$ and $\tilde{\alpha}$ are integral curves passing through the same point at time $t_1$ shows that they are both solutions of the same ODE, and hence they are equal in this chart (by the Picard-Lindelöf Theorem 0.22). This shows that, for any $t_1 \in \mathcal{T}$, there is an open set of times containing $t_1$ at which the two curves agree, meaning that there is an open neighborhood of $t_1$ contained in $\mathcal{T}$. Hence, $\mathcal{T}$ is open. As $(a, b)$ is connected, it follows that the clopen set $\mathcal{T} \subseteq (a, b)$ is equal to $(a, b)$.

Hence, we have shown that any two integral curves of $X$ that agree at one point must agree on their common domain. Now, for any $p \in M$, define $\mathcal{D}^p$ to be the union of all open intervals containing $0$ on which there is an integral curve of $X$ starting at $p$. By the above, for any $t \in \mathcal{D}^p$, all integral curves starting at $p$ have the same value at $t$; so we may define $\alpha_p(t)$ to be this common value. By definition, $\alpha_p$ is an integral curve defined on this union of intervals. It is maximal: if it could be extended, then that extension would have been included in the union. It is unique: any other integral curve agrees with $\alpha_p$ at the time $0$, and by the above, that means they agree everywhere. This concludes the proof. □

Proof of Theorem 5.15 For each $p \in M$, let $\mathcal{D}^p$ be the domain of the maximal integral curve of $X$ starting at $p$, and define $\mathcal{D} = \{(t, p) \in \mathbb{R} \times M : t \in \mathcal{D}^p \}$. We then define $\theta : \mathcal{D} \to M$ by $\theta(t, p) = \alpha_p(t)$. As usual, we will write $\theta^p(t) = \theta(t, p)$. So, in particular, $\theta^p = \alpha_p$, which gives part (a) by definition. We will verify the following things.

1. For any $p \in M$, if $s \in \mathcal{D}^p$ and $t \in \mathcal{D}^{\theta(s,p)}$ satisfy $s + t \in \mathcal{D}^p$, then $\theta(t, \theta(s, p)) = \theta(t + s, p)$.
2. Part (b).
3. $\mathcal{D}$ is open, and $\theta$ is smooth.
4. Part (c).

1. Fix $p \in M$, let $s \in \mathcal{D}^p$, and set $q = \theta(s, p) = \theta^p(s)$. The curve $\alpha : \mathcal{D}^p - s \to M$ defined by $\alpha = \tau_q \theta^p$ is an integral curve (by the translation lemma) and starts at $q$. By Lemma 5.16, $\alpha$ agrees
with \(\theta^q\) on their common domain, which is precisely to say that if \(t \in \mathcal{D}^{\theta(s,p)}\) with \(s + t \in \mathcal{D}^p\) then 
\[
\theta^q(t) = \tau_s \theta^p(t) = \theta^p(t + s). \quad \text{i.e.,} \quad \theta(t, \theta(s, p)) = \theta(s + t, p),
\]
as required.

(2) In part (1), we saw that the domain of \(\alpha\) was \(\mathcal{D}^p - s\); this must be contained in the domain of the maximal integral curve, so \(\mathcal{D}^p - s \subseteq \mathcal{D}^q = \mathcal{D}^{\theta(s,p)}\). On the other hand, since \(0 \in \mathcal{D}^p\), and so \(-s = 0 - s \in \mathcal{D}^p - s\), we have \(-s \in \mathcal{D}^q\), and by the group law (verified in part (1)) we get 
\[
\theta^q(-s) = p. \quad \text{Thus, applying the same argument with } (-s, q) \text{ in the place of } (s, p), \text{ we see that}
\]
\(\mathcal{D}^q + s \subseteq \mathcal{D}^p\), proving the reverse containment, and giving part (b).

(3) This is the tricky one. We define a set \(W \subseteq \mathcal{D}\) as follows: \((t, p) \in W\) iff there is some product neighborhood \(J \times U\), where \(J\) is an open interval containing \(0\) and \(t\), and \(U\) an open neighborhood of \(p\) in \(M\), on which \(\theta\) is defined and smooth. Note that, with \(t = 0\), if we take a chart at \(p\), then in local coordinates we may apply the Picard-Lindelöf theorem (part 3: smooth dependence on initial conditions) to conclude that \(\theta\) is defined an smooth on a product neighborhood of \((0, p)\).

Thus, \(W\) contains \(\{0\} \times M\), and so is not empty. What’s more, if \((t, p) \in W\), then all points in the neighborhood \(J \times U\) presumed to exist are also in \(W\) (they all have neighborhood \(J \times U\), and so \(W\) is open. We will show that \(W = \mathcal{D}\); this will show that \(\mathcal{D}\) is open, and (by definition of \(W\)) that \(\theta\) is smooth on its domain. We prove this by contradiction: suppose there is some point \((\tau, p_0) \in \mathcal{D} \setminus W\). By the above, \(\tau \neq 0\). Wlog, assume \(\tau > 0\) (the \(\tau < 0\) case is similar).

Let \(t_0 = \inf \{ t \in \mathbb{R} : (t, p_0) \notin W \}\). Since \(\theta\) is defined and smooth in some product neighborhood of \((0, p_0)\), we must have \(t_0 > 0\). Since \(t_0 \leq \tau\) and \(\mathcal{D}^{p_0}\) is an open interval containing \(0\) and \(\tau\), it follows that \(t_0 \in \mathcal{D}^{p_0}\). Set \(q_0 = \theta^{p_0}(t_0)\). By the above discussion, there is some neighborhood \(U_0\) of \(q_0\) and an interval \((-\epsilon, \epsilon)\) such that \((-\epsilon, \epsilon) \times U_0 \subseteq W\). The idea is to now use the group law to show that this neighborhood can be translated to give a smooth extension of \(\theta\) to a neighborhood of \((t_0, p_0)\), contradicting the definition of \(t_0\).

To be precise: choose some \(t_1 < t_0\) within distance \(\epsilon\) (so \(t_1 + \epsilon > t_0\)) and such that \(\theta^{p_0}(t_1) \in U_0\) (possible because \(\theta^{p_0}(t_0) = q_0, U_0\) is a neighborhood of \(q_0\), and \(\theta^{p_0}\) is continuous). Now, by definition of \(t_0\), since \(t_1 < t_0\) we have \((t_1, p_0) \in W\). This means there is a product neighborhood \((t_1 - t_1 +) \times U_1\) contained in \(W\), which means that \(\theta\) is defined and smooth there – and (since \(t_1 - t_0\)) also below: \(\theta\) is smooth on \([0, t_1 +) \times U_1\). By taking \(U_1\) small, we may assume that \(\theta(t_1, U_1) \subseteq U_0\) (since \(\theta(t_1, p_0) \in U_0\)). Now we define our extension: \(\tilde{\theta}: [0, t_1 + \epsilon) \times U_1 \to M\) is defined to be

\[
\tilde{\theta}(t, p) = \begin{cases} 
\theta_t(p), & p \in U_1, u \leq t < t_1, \\
\theta_{t-t_1} \circ \theta_{t_1}(p), & p \in U_1, t_1 - \epsilon < t < t_1 + \epsilon.
\end{cases}
\]

The group law for \(\theta\) guarantees that this is well-defined (it agrees on the overlap), and by the preceding discussion \(\tilde{\theta}\) is smooth. Now, by the translation lemma, each curve \(t \mapsto \tilde{\theta}(t, p)\) is an integral curve, and so \(\tilde{\theta}\) is a smooth extension of \(\theta\) to a neighborhood of \((t_0, p_0)\). This contradicts the choice of \(t_0\). Hence, there cannot have been any point \((\tau, p_0) \in \mathcal{D} \setminus W\), concluding the proof that \(W = \mathcal{D}\).

(4) Finally, we prove (c). Since \(\mathcal{D}\) is open, \(M_t = \{ p \in M: (t, p) \in \mathcal{D} \}\) is open (by the definition of the product topology on \(\mathbb{R} \times M\)). Using part (b), we have

\[
p \in M_t \implies t \in \mathcal{D}^p \implies \mathcal{D}^{\theta_t(p)} = \mathcal{D}^p - t \implies -t \in \mathcal{D}^{\theta_t(p)} \implies \theta_t(p) \in M_{-t}.
\]

This shows that \(\theta_t(M_t) \subseteq M_{-t}\). Now, the group law shows that \(\theta_t \circ \theta_{-t}(p) = p\) for all \(p \in M_{-t}\). What’s more, the same argument as above with \(t\) and \(-t\) reversed shows that \(\theta_{-t}(M_{-t}) \subseteq M_t\). So any \(p \in M_{-t}\) is equal to \(\theta_t(q)\) for a point \(q = \theta_{-t}(p) \in M_t\). This shows \(\theta_t\) maps \(M_t\) onto \(M_{-t}\).
The argument also shows that it is a bijection with inverse $\theta_{-t}$. Since the flow is smooth, both $\theta_t$ and $\theta_{-t}$ are smooth, and hence they are diffeomorphisms. \hfill \square

Thus, every smooth vector field $X$ is the infinitesimal generator of a unique maximal flow $\theta$; we call it the flow generated by $X$ or simply the flow of $X$.

To conclude our present discussion, let’s note that Lemma 5.8, which says that integral curves are natural, implies that the collection of all integral curves together (the flow) is also a natural construction.

**Lemma 5.17 ("flows are natural").** Let $M$ and $N$ be smooth manifolds, and let $F: M \to N$ be a smooth map. Let $X \in \mathcal{X}(M)$ and $Y \in \mathcal{X}(N)$, with flows $\theta$ and $\eta$ respectively. If $X, Y$ are $F$-related, then for each $t \in \mathbb{R}$ $F(M_t) \subseteq N_t$, and $\eta \circ F = F \circ \theta$ on $M_t$:

$$
\begin{array}{ccc}
M_t & \xrightarrow{F} & N_t \\
\theta_t & \downarrow & \eta \\
M_{-t} & \xrightarrow{F} & N_{-t}
\end{array}
$$

**Proof.** By Lemma 5.8, given an $p \in M$, the curve $F \circ \theta^p$ is an integral curve of $Y$, which starts at $F \circ \theta^p(p) = F(p)$. By uniqueness of (maximal) integral curves, the maximal integral curve $\eta^F(p)$ is defined on at least the interval $\mathcal{D}^p$, and $F \circ \theta^p = \eta^F(p)$ on this interval. This means that

$$
p \in M_t \implies t \in \mathcal{D}^p \implies t \in \mathcal{D}^F(p) \implies F(p) \in N_t,
$$

which is equivalent to $F(M_t) \subseteq N_t$. What’s more, the statement $F(\theta^p(t)) = \eta^F(p)(t)$ for all $t \in \mathcal{D}^p$ is precisely the statement that $\eta_t \circ F(p) = F \circ \theta_t(p)$ for all $p \in M_t$, as desired. \hfill \square

**Corollary 5.18.** Let $M$ and $N$ be a smooth manifolds, and let $F: M \to N$ be a diffeomorphism. If $X \in \mathcal{X}(M)$ with flow $\theta$, then the flow of $F_*(X)$ is $\eta_t = F \circ \theta_t \circ F^{-1}$, with domain $N_t = F(M_t)$ for each $t \in \mathbb{R}$.

As we mentioned just above Definition 5.12, a vector field is called complete if it is the infinitesimal generator of a global smooth flow. In light of Theorem 5.15, we can phrase this more accurately as: a vector field is complete if it’s flow is global, i.e. if its integral curves all exists for all time. It is generally quite hard to decide when this is the case. Example like 5.1 and 5.2 that can be solved completely explicitly are rare. Usually it is difficult to decide whether a given local integral curve can be extended. We do have the following tool.

**Lemma 5.19 (Uniform Time Lemma).** Let $M$ be a smooth manifold and $X \in \mathcal{X}(M)$. Suppose there is a uniform $\epsilon > 0$ so that, for every $p \in M$, the integral curve $\theta^p$ of $X$ starting at $p$ is defined on $(-\epsilon, \epsilon)$. Then $X$ is complete.

This is a typical “bootstrapping” argument. Let $q = \theta^p(\epsilon/2)$; by assumption $\theta^q$ also extends past time $\epsilon/2$, so let $r = \theta^q(\epsilon/2)$; by assumption, $\theta^r$ extends past time $\epsilon/2$, so let $s = \theta^r(\epsilon/2)$; and so forth. Then the group law shows that $\theta^p(t)$ exists up to time $3\epsilon/2$: $\theta^p(t) = \theta^s(t - \epsilon)$ for $\epsilon \leq t < 3\epsilon/2$. We can continue this way extending $\epsilon/2$ time units at a time, out to $\infty$. In the proof, we give a shorter treatment of this idea by a contradiction proof.

**Proof.** For a contradiction, we assume that there exists a $p \in M$ for which $\mathcal{D}^p$ is bounded above (a similar proof works in the case it is assumed to be bounded below). Fix some time $t_0$ with
\[
\sup \mathcal{D}^p - \epsilon < t_0 < \sup \mathcal{D}^p. \text{ Let } q = \theta^p(t_0). \text{ By assumption, } \theta^q \text{ is defined on } (-\epsilon, \epsilon). \text{ So we may define a curve } \alpha: (-\epsilon, t_0 + \epsilon) \text{ by }
\]
\[
\alpha(t) = \begin{cases}
\theta^p(t), & -\epsilon < t < \sup \mathcal{D}^p, \\
\theta^q(t-t_0), & t - \epsilon < t < t_0 + \epsilon.
\end{cases}
\]

By the group law, we have \( \theta^q(t - t_0) = \theta(t - t_0, q) = \theta(t - t_0, \theta^p(t_0)) = \theta_{t-t_0} \circ \theta_{t_0}(p) = \theta_t(p) = \theta^q(t) \). This shows that the two pieces of the definition of \( \alpha \) agree on their overlap. By the translation lemma, \( \alpha \) is an integral curve, and it starts at \( p \). But \( \alpha \) is defined on a domain strictly larger than the maximal domain \( \mathcal{D}^p \) of \( \theta^p \), which is a contradiction. \( \square \)

This gives us at least one class of vector fields that are complete.

**Theorem 5.20.** Every compactly supported smooth vector field is complete.

**Proof.** Let \( X \in \mathcal{X}(M) \) be compactly supported, with \( K = \text{supp } X \). Let \( p \in K \); then, since the flow \( \theta \) of \( X \) is defined on an open set, there is some product neighborhood \( (-\epsilon_p, \epsilon_p) \times U_p \) of \( (0, p) \) where \( \theta \) is defined. By compactness of \( K \), there are finitely many points \( p_1, \ldots, p_k \in K \) so that the open sets \( U_{p_1}, \ldots, U_{p_k} \) cover \( K \). Set \( \epsilon = \min\{\epsilon_{p_1}, \ldots, \epsilon_{p_k}\} \). Now, for any point \( q \in K \), choose some \( p_j \) such that \( q \in U_{p_j} \); then, since \( \theta \) is defined on \( (-\epsilon_{p_j}, \epsilon_{p_j}) \times U_{p_j} \), it follows that the integral curve \( \theta^q \) is defined on \( (-\epsilon_{p_j}, \epsilon_{p_j}) \), and so there exists an interval \( (-\epsilon, \epsilon) \). Hence, every integral curve starting in \( K \) is defined on \( (-\epsilon, \epsilon) \). If, on the other hand, \( q \notin K = \text{supp } X \), then \( X(q) = 0 \), and so the integral curve \( \theta^q \) is constant and defined for all time (so in particular on \( (-\epsilon, \epsilon) \)). It follows from the uniform time lemma that \( X \) is complete. \( \square \)

**Corollary 5.21.** Every smooth vector field on a compact smooth manifold is complete.

### 5.3. Lie Derivatives

Consider an old-fashioned vector field \( X: U \subseteq \mathbb{R}^n \to \mathbb{R}^n \). We were able to take derivatives of such vector fields by treating them as functions: the directional derivative in some direction \( v \) was, as usual
\[
D_vX(p) = \left. \frac{d}{dt} \right|_{t=0} X(p + tv) = \lim_{t \to 0} \frac{X(p + tv) - X(p)}{t}.
\]

Indeed, this simply means taking directional derivatives of the components independently. However, this heavily depends on the vector space structure of \( \mathbb{R}^n \) (so that we can take \( p + tv \)).

We could try to fix this as follows. Let \( X \in \mathcal{X}(M) \), and let \( p \in M \). For a vector \( v \in T_pM \), fix some curve \( \alpha: (-\epsilon, \epsilon) \to M \) so that \( \alpha(0) = p \) and \( \dot{\alpha}(0) = p \). Then we could try to define
\[
D_vX(p) = \lim_{t \to 0} \frac{X(\alpha(t)) - X(p)}{t}.
\]

But this is still problematic. First, which curve should we choose? There is no reason to think that the answer will depend only on \( \dot{\alpha}(0) \). But the bigger problem is: the difference quotient actually doesn’t make sense: \( X(\alpha(t)) \in T_{\alpha(t)}M \) while \( X(p) \in T_pM \), different vector spaces. If \( M = \mathbb{R}^n \), we can canonically identify these two spaces, but in general, there is no coordinate-free way to do so.

There is no solution to this problem: there is no way to take the directional derivative at some point \( p \) of a vector field in the direction of some vector living only in \( T_pM \) that is coordinate independent (i.e. well-defined). But there is a solution if we take the directional derivative in the direction of another vector field: the idea being that we should evaluate \( X \) at points along the flow of the other vector field in the difference quotient. Here is the definition.
Definition 5.22. Let $M$ be a smooth manifold, and fix two vector fields $X,Y \in \mathfrak{X}(M)$. Let $\theta$ be the flow of $Y$. The Lie derivative of $X$ with respect to $Y$ is defined to be the (rough) vector field $\mathcal{L}_Y(X)$ defined by

$$\mathcal{L}_Y(X)(p) = \left. \frac{d}{dt} \right|_{t=0} d(\theta_{-t})_{\theta_t(p)}(X_{\theta_t(p)}) = \frac{\lim_{t \to 0} d(\theta_{-t})_{\theta_t(p)}(X_{\theta_t(p)})}{t} - X_p. \quad (5.4)$$

Actually, we need to verify that this definition makes sense (i.e. that the limit exists), but at least the difference quotient itself makes sense. We evaluate $X$ at the point $\theta_t(p)$ that $p$ flows to under the flow of $X$. This gives us a vector in $T_{\theta_t(p)}M$, which we cannot compare to $X(p)$. So we transform that vector back into $T_pM$ by using the backwards flow, or rather the differential of the backwards flow, which is its infinitesimal action on vectors, giving us a vector back in $T_pM$ (since $\theta_{-t}: \theta_t(p) \mapsto p$). Note: although $\theta$ is not generally globally defined, for fixed $p$ the flow domain $\mathcal{D}_p$ is an interval containing 0, and so for small enough $t$ the difference quotient makes sense.

By Theorem 5.15 $\theta_{-t}$ is a diffeomorphism from $M_{-t} \to M_t$, where $M_{-t} = \{ q \in M : -t \notin \mathcal{D}^q \}$ is open. If we choose $t$ small enough that $-t \in \mathcal{D}_p$, then $p \in M_{-t}$; then we can restrict $X$ to the neighborhood $M_{-t}$ of $p$, and we have (from Proposition 4.16)

$$d(\theta_{-t})_{\theta_t(p)}(X(\theta_t(p))) = (\theta_{-t})_* (X)(p).$$

Thus, we can write

$$\mathcal{L}_Y(X)(p) = \left. \frac{d}{dt} \right|_{t=0} (\theta_{-t})_* (X)(p), \quad (5.5)$$

again, provided this limit makes sense. Note: we state this at a given point $p \in M$ to emphasize the point that the derivative being taken is of the function $t \mapsto (\theta_{-t})_* (X)(p)$, which is a map from some time interval into the tangent space $T_pM$: it is a regular calculus function.

Proposition 5.24 below shows that the Lie derivative does exist, and is in fact a familiar object. First we need the following lemma.

Lemma 5.23. Let $M$ be a smooth manifold, let $\epsilon > 0$, and let $f \in C^\infty((-\epsilon, \epsilon) \times M)$. For each $t \in (-\epsilon, \epsilon)$, let $f_t(p) = f(t,p)$. If $X \in C^\infty(M)$, then $(t,p) \mapsto X(f_t)$ is smooth function on $(-\epsilon, \epsilon) \times M$.

Proof. Let $(U, \varphi)$ be a chart in $M$, with coordinate functions $\varphi = (x^1, \ldots, x^n)$. Then we can write the vector field $X$ restricted to $U$ in the coordinate basis: for any $p \in U$,

$$X_p = \sum_{j=1}^n X^j(p) \left. \frac{\partial}{\partial x^j} \right|_p.$$

Write $f_t$ in local coordinates $\hat{f}_t = f_t \circ \varphi^{-1}$. Since $\varphi$ is a diffeomorphism $U \to \hat{U} = \varphi(U)$, the composition $(-\epsilon, \epsilon) \times \hat{U} \ni (t,x) \mapsto \hat{f}_t = f(t, \varphi^{-1}(X))$ is $C^\infty$, and therefore so are all its partial derivatives. Letting $\hat{p} = \varphi(p)$, we have

$$X(f_t)(p) = \sum_{j=1}^n X^j(p) \left. \frac{\partial \hat{f}_t}{\partial x^j} \right|_{\hat{p}},$$

which is therefore $C^\infty$ in both variables. As this holds true in any chart $U$, it holds globally.  \qed

Proposition 5.24. For any $X,Y \in \mathfrak{X}(M)$, let $\theta$ be the flow of $Y$. Then the function $t \mapsto (\theta_{-t})_* (X)(p)$ is smooth for each $p \in M$, and its derivative at $t = 0$ is $\mathcal{L}_Y(X) = [Y, X]$. 

Proof. We apply \((\theta_{-t})_*(X)|_p\) to a function \(f \in C^\infty(M)\); this gives
\[
(\theta_{-t})_*(X)f(p) = d(\theta_{-t})_{\theta_t(p)}(X_{\theta_t(p)})f = X_{\theta_t(p)}(f \circ \theta_{-t}).
\]
So we are supposed to take the limit as \(t \to 0\) of the difference quotient
\[
\frac{(\theta_{-t})_*(X)_p - X_p(f)}{t} = \frac{1}{t} \left[ X_{\theta_t(p)}(f \circ \theta_{-t}) - X_p f \right].
\]
Now we play the usual trick of adding and subtracting a connecting term: the above is equal to
\[
\frac{1}{t} \left[ X_{\theta_t(p)}(f \circ \theta_{-t}) - X_{\theta_t(p)}(f) + X_{\theta_t(p)}(f) - X_p f \right]
\]
The last two terms gives us \(YX|_p(f)\) in the limit as follows:
\[
\lim_{t \to 0} \frac{X_{\theta_t(p)}(f) - X_p f}{t} = \left. \frac{d}{dt} \right|_0 Xf(\theta_t(p)) = \left. \frac{d}{dt} \right|_0 (Xf) \circ \theta^p(t) = \dot{\theta}^p(0)(Xf) = Y_p(Xf) = YX|_p(f)
\]
since \(\theta^p\) is the maximal integral curve of \(Y\) that starts at \(p\) (so in particular \(\dot{\theta}^p(0) = Y(\theta^0(0)) = Y(p)\)). So, we have shown that
\[
\mathcal{L}_Y(X)|_p(f) = \lim_{t \to 0} \frac{X_{\theta_t(p)}(f \circ \theta_{-t}) - X_{\theta_t(p)}(f)}{t} + YX|_p(f)
\]
provided this limit exists.

The key here is Taylor’s theorem. For \(q \in M\) and \(t \in -\mathcal{D}^q\), let \(g(t, q) = f(\theta_{-t}(q))\); for fixed \(q \in M\), this is a smooth function \(-\mathcal{D}^q \to \mathbb{R}\). In particular, it is a smooth calculus function defined on some neighborhood of 0, and so Taylor’s theorem gives us
\[
g(t, q) = g(0, q) + t\vartheta_1 g(0, q) + O(t^2) = f(q) + t\vartheta_1 g(0, q) + O(t^2)\]
(since \(f(0, q) = f(\theta_0(q)) = f(q)\)). We’ll need more precise control over the \(g\)-dependence of the error term, so we use Theorem \[0.9\] which gives the error term as an integral. In fact, we need only expand to first order:
\[
f(\theta_{-t}(q)) = g(t, q) = f(q) + t \int_0^1 \vartheta_1 g(ts, q) \, ds \equiv f(q) + t h_1(q).
\]
As \(g\) is \(C^\infty\) in both variables in a neighborhood of \((0, p)\), the same is true of its partial derivatives and their integrals, so \(h_1\) is a smooth function (for all small enough \(t\) for which it is defined). We may apply the vector field
\[
X_{\theta_t(p)}(f \circ \theta_{-t}) = X_{\theta_t(p)}(f) + t X_{\theta_t(p)}(h_t).
\]
Hence, the remaining limit difference quotient in \((5.6)\) simplifies:
\[
\lim_{t \to 0} \frac{X_{\theta_t(p)}(f \circ \theta_{-t}) - X_{\theta_t(p)}(f)}{t} = \lim_{t \to 0} X_{\theta_t(p)}(h_t) = \lim_{t \to 0} (X(h_t))(\theta_t(p)).
\]
By Lemma \[5.23\], \((t, p) \mapsto X(h_t)(p)\) is a smooth function, and since \(\theta_t(p)\) is also a smooth function of \((t, p)\), this limit is equal to \(X(h_0)(\theta_0(p)) = X_p(h_0)\). But
\[
h_0(p) = \int_0^1 \vartheta_1 g(0, p) \, ds = \vartheta_1 g(0, p) = \left. \frac{\partial}{\partial t} \right|_{t=0} f(\theta_{-t}(p)).
\]
Applying the chain rule to the change of variables \(s = -t\) gives
\[
h_0(p) = -\left. \frac{\partial}{\partial s} \right|_{s=0} f(\theta_s(p)) = -\left. \frac{\partial}{\partial s} \right|_{s=0} f(\theta^p(s)) = -\dot{\theta}^p(0)(f) = -Y_p(f).
\]
Hence, finally combining with (5.6), we have
\[ \mathcal{L}_Y (X) \big|_p (f) = X(-Y(f)) \big|_p + YX(f) \big|_p = [Y, X](f) \big|_p. \]
This concludes the proof. \[ \square \]

Remark 5.25. The Lie bracket \([Y, X]\) and the Lie derivative \(\mathcal{L}_Y (X)\) were both known to be coordinate independent for a long time, but the only known proof of their equality required computations in local coordinates (which is the way \([1]\) approaches the computation). The above invariant proof is relatively new (less than half a century old).

While we are on the subject of Lie derivatives, we can give a similar definition of the Lie derivative of a smooth function with respect to a vector field: the result \(\mathcal{L}_X f\) is a new smooth function,
\[ \mathcal{L}_X f(p) = \frac{d}{dt} \bigg|_{t=0} f \circ \theta_t(p) = \lim_{t \to 0} \frac{f \circ \theta_t(p) - f(p)}{t}. \]
This, again, is an appealing interpretation of what should replace \(D_v f(p) = \frac{d}{dt} \bigg|_{t=0} f(p + tv)\) in the linear case, if \(v = X(p)\). Here we can see easily that
\[ \mathcal{L}_X f(p) = \frac{d}{dt} \bigg|_{t=0} f(\theta^p(t)) = \dot{\theta}^p(0)(f) = X(f). \]
This is consistent with our reinterpretation of tangent vectors as “directional derivative operators”.

So, in summary, we have Lie derivatives (so far) of functions and vector fields:
\[ \mathcal{L}_X f = Xf, \quad \mathcal{L}_X Y = [X, Y]. \]

The formula for the Lie derivative (being simply a Lie bracket) gives rise to a whole host of properties that are not at all obvious from the definition. We summarize them here.

**Proposition 5.26.** Let \(M\) be a smooth manifold, and let \(X, Y, Z \in \mathfrak{X}(M)\).

(a) \(\mathcal{L}_X Y = -\mathcal{L}_Y X\).
(b) \(\mathcal{L}_X [Y, Z] = [\mathcal{L}_X Y, Z] + [Y, \mathcal{L}_X Z]\).
(c) \(\mathcal{L}_{[X,Y]}(z) = \mathcal{L}_X \mathcal{L}_Y z - \mathcal{L}_Y \mathcal{L}_X z\).
(d) If \(g \in C^\infty(M)\), then \(\mathcal{L}_X(gY) = \mathcal{L}_X g \cdot Y + g \mathcal{L}_X Y = (Xg)Y + g \mathcal{L}_X Y\).
(e) If \(F: M \to N\) is a diffeomorphism, then \(F^\ast(\mathcal{L}_X Y) = \mathcal{L}_{F^\ast X}(F^\ast Y)\).

These are easy calculations left to the reader.

Remark 5.27. We already saw Proposition 5.26(c) show up as the Jacobi identity of Proposition 4.21(c); now we see that it really makes sense as a statement that the Lie derivative is a Lie derivation. Similarly, the mysterious identity of Proposition 4.21(d) is just Proposition 5.26(d) applied to \([fX, gY] = \mathcal{L}_{fX}(gY) = -\mathcal{L}_{gY}(fX)\) in each variable separately. Item (e) shows that the Lie derivative is a natural construction.

5.4. Commuting Vector Fields. We say two vector fields \(X, Y \in \mathfrak{X}(M)\) commute if \([X, Y] = 0\). In light of the previous section, this means that \(\mathcal{L}_X(Y) = 0\) (and also \(\mathcal{L}_Y(X) = 0\), so that “\(X\) does not vary in the \(Y\) direction” and vice versa. This is a little cumbersome to understand since the Lie derivative is a complicated object. In fact, it means the following.
Definition 5.28. Let \( X \in \mathfrak{X}(M) \), and let \( \theta \) be a smooth flow on \( M \). Say that \( X \) is invariant under the flow \( \theta \) if \( X \) is \( \theta_t \)-related to itself for each \( t \); more precisely, if \( X|_{M_t} \) is \( \theta_t \)-related to \( X|_{M_{t-\varepsilon}} \) for each \( t \in \mathbb{R} \). Equivalently, this says that

\[
d(\theta_t)_p(X_p) = X_{\theta_t(p)}, \quad (t, p) \in \mathcal{D}(\theta).
\]

So, to say \( X \) is invariant under \( \theta \) means that pushing it forward by the diffeomorphism \( \theta_t \) doesn’t change \( X \) for any \( t \). For example, it is clear that \( X \) is invariant under its own flow (this is precisely what it means for \( \theta^p \) to be an integral curve of \( X \)). In fact, \( X \) is invariant under the flow of any vector field it commutes with.

Theorem 5.29. Let \( X, Y \in \mathfrak{X}(M) \). TFAE:

(a) \( [X, Y] = 0 \).
(b) \( X \) is invariant under the flow of \( Y \).
(c) \( Y \) is invariant under the flow of \( X \).

Proof. We will show that (a) \( \iff \) (b); the equivalence (a) \( \iff \) (c) is the same (up to a minus sign). Let \( \theta \) be the flow of \( Y \).

(b) \( \implies \) (a): By assumption, \( X_{\theta_t(p)} = d(\theta_t)_p(X_p) \) whenever \( (t, p) \in \mathcal{D}(\theta) \). That is: \( X_p = d(\theta_{-t})_{\theta_t(p)}(X_{\theta_t(p)}) \) (since \( \theta_{-t} = \theta_t^{-1} \)). Referring to the (5.4) defining the Lie derivative, it follows immediately that \( \mathcal{L}_X(Y)|_p = 0 \). By Proposition 5.24 it follows that \( [X, Y]|_p = 0 \). Since every \( p \in M \) is in \( M_t \) and \( M_{t-\varepsilon} \) for all sufficiently small \( t \), this shows \( [X, Y] = 0 \), confirming (a).

(a) \( \implies \) (b): We are assuming that \( \mathcal{L}_X(Y) = [Y, X] = 0 \). Fix \( p \in M \), and define a function \( \alpha: \mathcal{D}^p \to T_p M \) by

\[
\alpha(t) = d(\theta_{-t})_{\theta_t(p)}(X_{\theta_t(p)}) = (\theta_{-t})_*(X)|_p.
\]

Then \( \alpha \) is differentiable: it is clearly smooth on \( \mathcal{D}^p \setminus \{0\} \), and by (5.5) its derivative at 0 is \( \mathcal{L}_Y(X)|_p = 0 \). In fact, let us compute \( \alpha'(t_0) \) for any \( t_0 \in \mathcal{D}^p \) (that is, the usual derivative of a curve in a vector space). We do this by translating \( t_0 \) to 0.

\[
\alpha'(t_0) = \frac{d}{ds} \bigg|_{s=0} \alpha(t_0 + s) = \frac{d}{ds} \bigg|_{s=0} d(\theta_{-t_0-s})_{\theta_t_0+s}(p).
\]

Now, \( \theta_{t_0+s} = \theta_{t_0} \circ \theta_s \) on a neighborhood of \( p \), and since \( \theta_{t_0} \) is independent of \( s \), we therefore have

\[
\alpha'(t_0) = d(\theta_{-t_0}) \left( \frac{d}{ds} \bigg|_{s=0} d(\theta_{-s})_{\theta_{t_0+s}(p)} \right) = d(\theta_{-t_0}) \left( \mathcal{L}_Y(X)|_{\theta_{t_0}(p)} \right) = 0.
\]

Hence, \( \alpha \) is constant, and since \( \alpha(0) = X_p \), we have \( \alpha(t) = X_p \) for all \( t \in \mathcal{D}^p \). This is precisely to say that \( X \) is invariant under \( \theta \).

Example 5.30. We can use Theorem 5.29 to characterize what vector fields are invariant under the flows of Examples 5.1, 5.2, and 5.11.

- **Translation:** the flow \( \tau_t(x, y) = (x + t, y) \) is generated by the vector field \( Y = \partial / \partial x \). Let \( X = X^1 \partial / \partial x + X^2 \partial / \partial y \) be a vector field on \( \mathbb{R}^2 \) that is invariant under translations in the \( e_1 \) direction, meaning invariant under the flow \( \tau \). This is the same as insisting that \( [X, Y] = 0 \).

We can compute, for any \( f \in C^\infty(\mathbb{R}^2) \),

\[
[X, Y](f) = \left( X^1 \frac{\partial}{\partial x} + X^2 \frac{\partial}{\partial y} \right) \frac{\partial f}{\partial x} - \frac{\partial}{\partial x} \left( X^1 \frac{\partial f}{\partial x} + X^2 \frac{\partial f}{\partial y} \right) = -\frac{\partial X^1}{\partial x} \frac{\partial f}{\partial x} - \frac{\partial X^2}{\partial x} \frac{\partial f}{\partial y}.
\]
In order for this to be 0 for all \( f \), it is necessary and sufficient that \( \frac{\partial X_1}{\partial x} = \frac{\partial X_1}{\partial x} = 0 \). Hence, for a vector field \( X \) to be invariant under translations in the \( x \)-direction, it is necessary and sufficient for its coefficients to be constant in \( x \).

- **Rotation:** the flow \( R_t(x, y) = (x \cos t - y \sin t, x \sin t + y \cos t) \) is generated by the vector field \( Y = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \). A vector field \( X \) is invariant under rotations (under the flow \( R \)) if and only if \([X, Y] = 0\). Here it will be convenient to work in polar coordinates \((r, \theta)\) (on the plane minus the negative \( x \)-axis, for example). Here \( Y = \frac{\partial}{\partial \theta} \), and so expanding \( X = X^1 \frac{\partial}{\partial r} + X^2 \frac{\partial}{\partial \theta} \), we have exactly the same calculation as above: \([X, Y] = 0 \) iff \( X^1 \) and \( X^2 \) are constant functions of \( \theta \). That is: a vector field is rotationally-invariant iff its components depend only on the radial variable.

- **Dilation:** the Euler vector field \( E = x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \) generates the flow \( \phi_t(p) = e^t p \). Here again, it will pay to convert to polar coordinates, where \( E = \frac{\partial}{\partial r} \). Again by the same argument as above, it follows that a vector field \( X \) is invariant under the dilation flow \( \phi_t \) if and only if \([E, X] = 0\), which means \( X \) (written in polar coordinates) has coefficients that only depend on \( \theta \), not on \( r \).

The clearest and deepest way to understand what it means for two vector fields to commute is to say that it means their flows commute: if \( \theta \) is the flow of \( X \) and \( \phi \) is the flow of \( Y \), then \([X, Y] = 0\) if and only if \( \theta_t \circ \phi_s = \phi_s \circ \theta_t \) for all \( s, t \). However, in general this is quite trickily to really make sense of: if \( X, Y \) are not complete, then the domains of the two flows may not be very compatible, and the naïve guess for what it really means for them to commute turns out to be wrong: it is possible for \([X, Y] = 0\) and yet \( \theta_t \circ \phi_s(p) \neq \phi_s \circ \theta_t(p) \) for some point \( p \) and times \( t, s \) for which both sides are defined. (The trouble is that \( \mathcal{D}^p(\theta) \) and \( \mathcal{D}^p(\phi) \) must be intervals containing 0, but when composing this recenters and one can be looking at the “wrong” integral curve – one that doesn’t go through 0.) It is quite annoying to state and prove the right general theorem; we will content ourselves with the statement for complete vector fields presently.

**Theorem 5.31.** Let \( X, Y \in \mathcal{X}(M) \) be complete vector fields. Then \([X, Y] = 0 \) iff their (global) flows \( \theta, \phi \) commute in the sense that \( \theta_t \circ \phi_s = \phi_s \circ \theta_t \) for all \( s, t \in \mathbb{R} \).

**Proof.** First, assume \([Y, X] = \mathcal{L}_Y(X) = 0 \). By Theorem 5.29, this means that \( X \) is invariant under \( \phi \), which means precisely that, for each \( s \in \mathbb{R} \), \( (\phi_s)_*(X) = X \). Now, by Corollary 5.18 the flow of \((\phi_s)_*(X) \) is \( \phi_s \circ \theta_t \circ \phi_s^{-1} \); but since \( (\phi_s)_*(X) = X \), this means that \( \phi_s \circ \theta_t \circ \phi_s^{-1} = \theta_t \). This shows that \( \theta \) and \( \phi \) commute.

Conversely, suppose the flows commute. This can be written in the form \( \phi^{\theta_t}(p)(s) = \theta_t(\phi^p(s)) \).

Now differentiate both sides with respect to \( s \) at \( s = 0 \). Since \( \phi^p(s) \) is an integral curve, the left-hand side is \( Y_{\theta_t}(p) \). The right-hand side becomes

\[
\left. \frac{d}{ds} \right|_{s=0} \theta_t(\phi^p(s)) = d(\theta_t)_p(\phi^p(0)) = d(\theta_t)_p(Y_p).
\]

So \( Y_{\theta_t}(p) = d(\theta_t)_p(Y_p) \), which means that \((\theta_t)_*(Y) = Y \) holds for all \( t \). Replacing \( t \) with \(-t\), this means that \( \mathcal{L}_Y(X) = -\frac{d}{dt}_{t=0} (\theta_{-t})_*(Y) = -\frac{d}{dt}_{t=0} Y = 0 \), as desired. \( \square \)
6. The Cotangent Bundle and 1-Forms

We have thus far failed to mention what are arguably the most important vector fields from classical vector calculus: conservative vector fields, which have the form \( \nabla f \) for some smooth function \( f \). The reason they have not come up yet is because — spoiler alert—they are not vector fields at all. To be clear about what this means, let’s take an example: let \( M = \mathbb{R}^2_+ \) be the right half-plane, \( \{(x, y) \in \mathbb{R}^2 : x > 0\} \). Let \( f \in C^\infty(M) \), and consider the vector field \( \nabla f \). In our present notation, this should be the section of \( TM \) whose expression in the (global) local coordinates \((x, y)\) is given by

\[
\nabla f|_p = \frac{\partial f}{\partial x}(\hat{p}) \frac{\partial}{\partial x}|_p + \frac{\partial f}{\partial y}(\hat{p}) \frac{\partial}{\partial y}|_p \tag{6.1}
\]

where \( \hat{p} = (x, y) \) is the coordinate expression for the point \( p \) in Cartesian coordinates. Now, if \( \nabla f \) is really a vector field on the manifold \( M \), this means (cf. Section 3.3) that we must have the same expression in any coordinates. For example, it must also be true that, using polar coordinates \((r, \theta)\) on \( M \) (which are also globally defined),

\[
\nabla f|_p = \frac{\partial f}{\partial r}(\hat{p}) \frac{\partial}{\partial r}|_p + \frac{\partial f}{\partial \theta}(\hat{p}) \frac{\partial}{\partial \theta}|_p \tag{6.2}
\]

where now \( \hat{p} \) denotes the coordinate expression for \( p \) in polar coordinates. But we can check if this is really the case, by using the transformation law (again cf. Section 3.3). In particular, we have

\[
\frac{\partial}{\partial x} = \frac{\partial r}{\partial x} \frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x} \frac{\partial}{\partial \theta} = \cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta}
\]

\[
\frac{\partial}{\partial y} = \frac{\partial r}{\partial y} \frac{\partial}{\partial r} + \frac{\partial \theta}{\partial y} \frac{\partial}{\partial \theta} = \sin \theta \frac{\partial}{\partial r} + \frac{\cos \theta}{r} \frac{\partial}{\partial \theta}
\]

Thus, from (6.1), we have

\[
\nabla f|_p = \frac{\partial f}{\partial x} \left( \cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \right) + \frac{\partial f}{\partial y} \left( \sin \theta \frac{\partial}{\partial r} + \frac{\cos \theta}{r} \frac{\partial}{\partial \theta} \right)
\]

\[
= \left( \cos \theta \frac{\partial f}{\partial x} + \sin \theta \frac{\partial f}{\partial y} \right) \frac{\partial}{\partial r} + \frac{1}{r} \left( -\sin \theta \frac{\partial f}{\partial x} + \cos \theta \frac{\partial f}{\partial y} \right) \frac{\partial}{\partial \theta}
\]

(where we implicitly write \( \frac{\partial}{\partial x} \) to mean \( \frac{\partial f}{\partial x}(r \cos \theta, r \sin \theta) \) here, and similarly with \( \frac{\partial}{\partial y} \)). Now we can also transform the derivatives in the coefficients to express them in terms of the vector fields \( \frac{\partial}{\partial r} \) and \( \frac{\partial}{\partial \theta} \), in the same manner:

\[
\cos \theta \frac{\partial f}{\partial x} + \sin \theta \frac{\partial f}{\partial y} = \cos \theta \left( \cos \theta \frac{\partial f}{\partial r} - \frac{\sin \theta}{r} \frac{\partial f}{\partial \theta} \right) + \sin \theta \left( \sin \theta \frac{\partial f}{\partial r} + \frac{\cos \theta}{r} \frac{\partial f}{\partial \theta} \right)
\]

\[
= \left( \cos^2 \theta + \sin^2 \theta \right) \frac{\partial f}{\partial r} + \frac{1}{r} \left( -\cos \theta \sin \theta + \sin \theta \cos \theta \right) \frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial r},
\]

while

\[
-\sin \theta \frac{\partial f}{\partial x} + \cos \theta \frac{\partial f}{\partial y} = -\sin \theta \left( \cos \theta \frac{\partial f}{\partial r} - \frac{\sin \theta}{r} \frac{\partial f}{\partial \theta} \right) + \cos \theta \left( \sin \theta \frac{\partial f}{\partial r} + \frac{\cos \theta}{r} \frac{\partial f}{\partial \theta} \right)
\]

\[
= \left( -\sin \theta \cos \theta + \cos \theta \sin \theta \right) \frac{\partial f}{\partial r} + \frac{1}{r} \left( \sin^2 \theta + \cos^2 \theta \right) \frac{\partial f}{\partial \theta} = \frac{1}{r} \frac{\partial f}{\partial \theta}.
\]
Now combining, this shows that the gradient field $\nabla f$, defined in Cartesian coordinates by (6.1), converted to polar coordinates, has the form

$$
\nabla f|_p = \frac{\partial f}{\partial r}(\hat{p}) \frac{\partial}{\partial r}|_p + \frac{1}{r^2} \frac{\partial f}{\partial \theta}(\hat{p}) \frac{\partial}{\partial \theta}|_p.
$$

This does not match (6.2) (there is a factor of $\frac{1}{r^2}$ discrepancy in the $\frac{\partial}{\partial \theta}$ term).

Thus, the gradient, viewed as a vector field in Cartesian coordinates, is most definitely coordinate dependent: it does not generalize to well-defined map from $M \rightarrow TM$. In older language, we would say it is not invariant, or coordinate-dependent. There was a strong clue that this would happen: in (6.1), there are $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ terms repeated; we know that, in our transformation laws (really just the chain rule), $\frac{\partial}{\partial x}$ should come with $\frac{\partial x}{\partial u}$ for some new coordinate (function) $u$; not with $\frac{\partial f}{\partial x}$.

This is not a failing, however. Instead, it introduces us to a new kind of vector field – one that transforms in the opposite way, with the $\frac{\partial}{\partial x}$ on the bottom – meaning that it transforms “in the same direction” as the derivatives. Such vector fields will be called covariant (transform the same direction) as opposed to contravariant (the vector fields we’ve been working with thus far are called contravariant in older literature). Note: these terms are unrelated to the words covariant and contravariant as they are used in category theory, so beware!

As we’ll see, covariant vector fields are actually more natural and fundamental in many respects.

6.1. Dual Spaces and Cotangent Vectors. Let $V$ be a finite-dimensional real vector space. The dual space $V^*$ is defined to be $V^* = L(V, \mathbb{R})$, the set of linear maps from $V$ to $\mathbb{R}$. This is a vector space in its own right, under the usual operations. It is, in fact, isomorphic to $V$. To see this, select a basis $\{e_1, \ldots, e_n\}$ for $V$. The associated dual vectors $\{e_1^*, \ldots, e_n^*\}$ in $V^*$ are defined by

$$
e_j^* (v^1 e_1 + \cdots + v^n e_n) = v^j.
$$

I.e. $e_j^*$ are the linear extensions of $e_j^*(e_i) = \delta^j_i$. These are certainly linear functionals on $V$. They form a basis:

- They are linearly independent: if $\sum_j \alpha_j e_j^* = 0$ (as a linear functional) this means that $\sum_j \alpha_j e_j^*(v) = 0$ for all $v \in V$; in particular this means that $0 = \sum_i \alpha_i e_i^* (e_i) = \alpha_i$ for each $i$, so $\alpha_1 = \cdots = \alpha_n = 0$.

- They span $V^*$: if $\lambda \in V^*$, define $\omega = \lambda(e_1) e_1^* + \cdots + \lambda(e_n) e_n^*$. We compute for any $v = \sum_j v^j e_j \in V$ that $\omega(v) = \sum_j \lambda(e_j) e_j^*(v) = \sum_j v^j \lambda(e_j) = \lambda(\sum_j v^j e_j) = \lambda(v)$. Thus $\lambda = \omega \in \text{span}\{e_1^*, \ldots, e_n^*\}$.

Thus, $\dim V^* = \dim V$, and so $V$ and $V^*$ are isomorphic as vector spaces. Indeed, we could define an isomorphism $V \rightarrow V^*$ to be the linear extension of $e_j \mapsto e_j^*$, $1 \leq j \leq n$.

There’s nothing wrong with this, but it is not natural, which we can make sense of simply here by saying that it is basis dependent. Indeed, define this isomorphism $\phi: V \rightarrow V^*$ by $\phi(e_j) = e_j^*$ for the given chosen basis. Let’s write this in terms of a different basis $\{f_1, \ldots, f_n\}$ for $V$. Since it is also a basis, there is an invertible $n \times n$ matrix $A$ so that $f_j = \sum_i A^i_j e_i$ for all $j$. Then we have

$$
\phi(f_j) = \sum_i A^i_j \phi(e_i) = \sum_i A^i_j e_i^*.
$$

We might hope that this is equal to $f_j^*$. To check this, we must express the $e_i$ in terms of the $f_j$, which means taking the inverse matrix $B = A^{-1}$, so that $e_j = \sum_i B^j_i f_i$. Then we can write $f_j^*$ in
the basis \( \{e_j^*\} \) as follows:

\[
f_j^*(v) = f_j^* \left( \sum_i v^i e_i \right) = f_j^* \left( \sum_i v^i \sum_k B_{ki}^j f_k \right) = \sum_{i,k} v^i B_{ki}^j f_k = \sum_i v^i B_i^j = \sum_i B_i^j e_i^*(v)
\]

which shows that

\[
f_j^* = \sum_i B_i^j e_i^* = \sum_i [B^\top]^i_j e_i^*
\]

where \( B^\top \) is the transpose of \( B \) (transposing rows and columns). Thus, we see that \( (6.3) \) does not in general simplify to \( \phi(f_j) = f_j^* \) – this is only the case when the change of basis matrix happens to satisfy \( A = B^\top = (A^{-1})^\top \).

**Remark 6.1.** Note this says that if we restrict our basis-changes to orthogonal transformations – i.e. where the change of basis matrix \( A \) is orthogonal, \( A^\top A = I \) – then this isomorphism is natural. In fact, if we fix an inner product \( \langle \cdot, \cdot \rangle \) on \( V \) to begin with, and insist that all our bases be orthonormal bases with respect to the inner product, then the change of basis matrix will indeed always be orthogonal. So we see that, in the category of finite-dimensional inner product spaces, \( V \) and \( V^* \) really are canonically isomorphic. It is easy to check that the isomorphism \( e_j \mapsto e_j^* \) in this restricted setting has the basis independent representation

\[
v \mapsto \langle \cdot, v \rangle.
\]

**Remark 6.2.** The inspired reader might see similarities between the above calculations and the one we did changing the gradient from Cartesian to polar coordinates. This is no accident, as this restricted setting has the basis independent representation

\[
\lambda \mapsto \lambda(\text{grad} f).
\]

Note that the transpose of the change-of-basis matrix came into play above. In fact, the transpose of a matrix is closely related to the dual space. Let \( T: V \to W \) be a linear map. Then there is an induced **dual map** \( T^*: W^* \to V^* \), which is simply defined by

\[
(T^*\lambda)(v) = \lambda(Av).
\]

It is easy to verify that, for any \( \lambda \in W^* \), \( T^*\lambda \in V^* \), so this is well-defined, and moreover \( T^* \) is a linear map. We record some key properties of it in the next lemma, which is left as an exercise to prove.

**Lemma 6.3.** Let \( V, W \) be finite-dimensional real vector spaces, and let \( S, T: V \to W \) be linear maps. Then we have the following.

(a) \((\text{Id}_V)^* = \text{Id}_{V^*}\).

(b) \((S \circ T)^* = T^* \circ S^*\).
Lemma 6.4. If \( \{v_j\} \) is a basis for \( V \) and \( \{w_j\} \) is a basis for \( W \), and \( A \) is the matrix of \( T \) in terms of these bases, then the matrix of \( A^\ast \) in terms of the dual bases \( \{v^*_j\} \) and \( \{w^*_j\} \) is \( A^\top \).

In fact, the dual map \( T^* \) is often called the transpose of \( T \). This is a nice way to see that transpose is a natural construction: it commutes with changing bases, because it is really given by a basis-free object.

Nevertheless, even though there is a natural transpose map, it is still the case (as we saw above) that the isomorphism \( e_j \mapsto e^*_j \) is basis dependent. In fact, it is a theorem that there does not exist a basis-independent isomorphism. (Also: the arguments above all fail if \( V \) is not finite-dimensional. Although the map \( e_j \mapsto e^*_j \) is still an injective linear map \( V \to V^\ast \) in that case, it is never surjective. But that won’t bother us presently.)

Now, we can repeat the process: having constructed \( V^\ast \), we may take its dual space \((V^\ast)^\ast = V^{\ast\ast}\), the space of linear functionals \( L(V^\ast, \mathbb{R}) \). Again, there will be no basis-independent isomorphism \( V^\ast \to V^{\ast\ast} \). Somewhat miraculously, though, we may compose two basis-dependent isomorphisms, and get a basis-independent one.

**Lemma 6.4.** The canonical map \( \xi : V \to V^{\ast\ast} \) defined by

\[
\xi v(\lambda) = \lambda(v)
\]

is an isomorphism.

**Proof.** Fix a basis \( \{e_j\} \) of \( V \), and its dual basis \( \{e^*_j\} \) of \( V^\ast \). Then we know \( \alpha : V \to V^\ast \) defined by \( \alpha(e_j) = e^*_j \) is an isomorphism. Similarly, we define \( \beta : V^\ast \to V^{\ast\ast} \) by \( \beta(e^*_j) = e^{\ast\ast}_j \) (using the dual basis to \( \{e^*_j\} \)), and this is also an isomorphism. Then \( \beta \circ \alpha : V \to V^{\ast\ast} \) is an isomorphism.

We claim that \( \xi = \beta \circ \alpha \), and is therefore an isomorphism.

To see this, first note that \( \xi \) is linear (by elementary computation), so it suffices to check that \( \xi \) and \( \beta \circ \alpha \) agree on a basis of \( V \); we will, of course, use the basis \( \{e_j\} \). On the one hand, we have for any \( \lambda \in V^\ast \)

\[
\xi(e_j)(\lambda) = \lambda(e_j).
\]

On the other hand,

\[
\beta \circ \alpha(e_j)(\lambda) = \beta(e^*_j)(\lambda) = e^{\ast\ast}_j(\lambda).
\]

Now expanding \( \lambda = \sum_i \lambda(e_i)e^*_i \), we therefore have

\[
\beta \circ \alpha(e_j)(\lambda) = e^{\ast\ast}_j \left( \sum_i \lambda(e_i)e^*_i \right) = \sum_i \lambda(e_i)\delta^*_j = \lambda(e_j).
\]

Thus \( \xi \) is an isomorphism. Although our proof used a basis, the definition of \( \xi \) is manifestly basis-independent. (Note: in the infinite-dimensional case, \( \xi \) is no longer an isomorphism, but it is still injective; it is called the canonical embedding of \( V \) into \( V^{\ast\ast} \).)

Now, let \( M \) be a smooth manifold, and let \( p \in M \). The cotangent space at \( p \), denoted \( T^*_p M \), is the dual to the tangent space:

\[
T^*_p M \equiv (T_p M)^\ast.
\]

We refer to elements of \( T^*_p M \) as **cotangent vectors**, and frequently use lower-case greek letters like \( \omega_p, \lambda_p \in T^*_p M \). Fix a chart \((U, \varphi)\) at \( p \), with coordinate functions \( \varphi = (x^1, \ldots, x^n) \). Then the coordinate vectors \( \left\{ \frac{\partial}{\partial x^i}|_p \right\}_{1 \leq i \leq n} \) form a basis for \( T_p M \). To avoid messy notation, for now let
us refer to the dual basis as \( \{ \lambda_j^i \}_{1 \leq j \leq n} \). From the proof above that the dual basis spans the dual space, we can then express any cotangent vector \( \omega_p \in T^*_p M \) as

\[
\omega_p = \sum_{j=1}^{n} \omega_p^j \lambda_j^i, \quad \text{where} \quad \omega_p^j = \omega_p \left( \frac{\partial}{\partial x^j} \bigg|_p \right).
\]

Now, what happens if we change coordinates? Let \((V, \psi)\) be another chart at \( p \) with coordinate functions \( \psi = (y^1, \ldots, y^n) \). Denote the dual basis to \( \{ \frac{\partial}{\partial y^i} \big|_p \}_{1 \leq i \leq n} \) as \( \{ \mu_j^i \}_{1 \leq j \leq n} \). We want to compute the components \( \tilde{\omega}_p^j \) of \( \tilde{\omega} \) in terms of this new dual basis. As above, these coefficients are given simply by

\[
\tilde{\omega}_p^j = \omega_p \left( \frac{\partial}{\partial y^j} \bigg|_p \right).
\]

To relate the \( \omega_p^j \) to the \( \tilde{\omega}_p^j \), first convert the coordinate vectors:

\[
\frac{\partial}{\partial x^j} \bigg|_p = \sum_{i=1}^{n} \frac{\partial y^i}{\partial x^j}(\hat{p}) \frac{\partial}{\partial y^i} \bigg|_p.
\]

Hence, we have

\[
\omega_p^j = \omega_p \left( \frac{\partial}{\partial x^j} \bigg|_p \right) = \omega_p \left( \sum_{i=1}^{n} \frac{\partial y^i}{\partial x^j}(\hat{p}) \frac{\partial}{\partial y^i} \bigg|_p \right) = \sum_{i=1}^{n} \frac{\partial y^i}{\partial x^j}(\hat{p}) \omega_p^i.
\]

Now, compare this to how the components of vector change under the same coordinate change: if \( T_p M \ni X_p = \sum_{j=1}^{n} X_p^j \frac{\partial}{\partial x^j} \big|_p \), then we have

\[
X_p = \sum_{j=1}^{n} X_p^j \left( \sum_{i=1}^{n} \frac{\partial y^i}{\partial x^j}(\hat{p}) \frac{\partial}{\partial y^i} \bigg|_p \right) = \sum_{i=1}^{n} \left( \sum_{j=1}^{n} \frac{\partial y^i}{\partial x^j}(\hat{p}) X_p^j \right) \frac{\partial}{\partial y^i} \bigg|_p
\]

which is to say that the components \( \tilde{X}_p^i \) of the vector \( X_p \) in terms of the new coordinate basis are given by

\[
\tilde{X}_p^i = \sum_{j=1}^{n} \frac{\partial y^i}{\partial x^j}(\hat{p}) X_p^j.
\]

Let us restate (6.4) and (6.5) for direct comparison in the following proposition.

**Proposition 6.5.** Let \( M \) be a smooth manifold and \( p \in M \). Fix two charts \((U, \varphi = (x^j)_{j=1}^{n})\) and \((V, \psi = (y^i)_{i=1}^{n})\) at \( p \). Let \( X_p \in T_p M \) and \( \omega_p \in T^*_p M \). If \( X \) has coordinates \( X_p^j \) in terms of the \( \varphi \) coordinate basis, and \( \omega_p \) has coordinates \( \omega_p^j \) in terms of its dual basis, then the components \( \tilde{X}_p^i \) of \( X_p \) in terms of the \( \psi \) coordinate basis, and the components \( \tilde{\omega}_p^j \) of \( \omega_p \) in terms of its dual basis, are given by

\[
\tilde{X}_p^i = \sum_{j=1}^{n} \frac{\partial y^i}{\partial x^j}(\hat{p}) X_p^j, \quad \tilde{\omega}_p^j = \sum_{i=1}^{n} \frac{\partial y^i}{\partial x^j}(\hat{p}) \omega_p^i.
\]

Thus, the two transform opposite to each other. Indeed, we could restate the \( \omega \) coordinate transformation as follows: the matrix \( \frac{\partial y^i}{\partial x^j}(\hat{p}) \) is just the Jacobian of the transition map \( \psi \circ \varphi^{-1} \) at the point \( \varphi(p) \). Its inverse is the Jacobian of the inverse map at \( \psi(p) \) (which we also, somewhat
confusingly in this case, denote \( \hat{p} \)), and so we can write the transformation law for the components of \( \omega_p \) in the form

\[
\tilde{\omega}_p^i = \sum_{j=1}^n \frac{\partial x^j}{\partial y^i}(\hat{p}) \omega_p^j.
\]

Now the \( \partial x^i \)'s are on top. This is the key to understanding the invariant generalization of a gradient.

6.2. The Differential, Reinterpreted. Let \( f \in C^\infty(M) \), fix \( p \in M \), and let \( q = f(p) \). The differential map \( df_p \) is a linear map \( T_p M \to T_q \mathbb{R} \). Of course, the tangent space \( T_q \mathbb{R} \) can be identified with \( \mathbb{R} \) itself, in a natural way. In terms of the naïve definition of \( T_q \mathbb{R} = \{(q,v) : v \in \mathbb{R}\} \), we simply identify \( (q,v) \cong v \). In our more invariant language, what does this mean? To see the answer, we work in local (global) coordinates: fix the usual Cartesian coordinate \( x \) on \( \mathbb{R} \), also known as the identity map \( \text{Id}_\mathbb{R} \). Then any element \( X_q \in T_q \mathbb{R} \) can be written uniquely as

\[ X_q = X_q^1 \frac{\partial}{\partial x} \bigg|_q \]

for some coefficient \( X_q^1 \in \mathbb{R} \). Of course, the identification we seek is \( X_q \cong X_q^1 \). What does this mean on the invariant level? Evidently,

\[ X_q^1 = X_q^1 \frac{\partial}{\partial x} \bigg|_q (x) = X_q(x) = X_q(\text{Id}_\mathbb{R}). \]

Because of the 1-dimensionality, the map \( T_q \mathbb{R} \to \mathbb{R} \) given by \( X \mapsto X(\text{Id}_\mathbb{R}) \) is an isomorphism. Hence, we can think of \( df_q \) as a map \( T_q M \to \mathbb{R} \), so long as we post compose with this isomorphism. This gives us a new interpretation of the differential, which for the moment we call \( \tilde{df}_p \):

\[
\tilde{df}_p(X_p) = df_p(X_p)(\text{Id}_\mathbb{R}) = X_p(\text{Id}_\mathbb{R} \circ f) = X_p(f).
\]

We now immediately drop the \( \sim \) and refer to this also as the differential, keeping in mind that, although it is a different object, it is the same as the old differential modulo the identification \( T_q \mathbb{R} \cong \mathbb{R} \). And note, what kind of object is it? \( df_p \) is a linear map \( T_p M \to \mathbb{R} \): it is a cotangent vector.

**Definition 6.6.** Let \( M \) be a smooth manifold, \( p \in M \), and \( f \in C^\infty(M) \). The **differential of \( f \) at \( p \)**, \( df_p \in T_p^* M \) is the cotangent vector defined by

\[ df_p(X_p) = X_p(f). \]

Note that, in our invariant language, \( T_p M \) consists of derivations at \( p \): a certain class of linear functionals on the vector space \( C^\infty(M) \). That is: \( T_p M \) is a subspace of \( C^\infty(M)^* \). Thus, the definition above is akin to the canonical embedding \( \xi : C^\infty(M) \to C^\infty(M)^{**} \). We have to be a little careful, since we are mapping \( C^\infty(M) \) to the dual space of a **subspace** of \( C^\infty(M)^* \); as such, it is not actually one-to-one.

In local coordinates \( (U, \varphi = (x^i)_{j=1}^n) \) at \( p \in M \), the coordinate functions \( x^j \) are smooth, and so they have differentials \( dx^j_p \). We can express them in terms of the dual basis to the coordinate vectors \( \left. \frac{\partial}{\partial x^i} \right|_p \):

\[
dx^j_p \left( \left. \frac{\partial}{\partial x^i} \right|_p \right) = \left. \frac{\partial}{\partial x^i} \right|_p (x^j) = \delta^j_i.
\]
This is precisely to say that
\[ dx^j_p = \left( \frac{\partial}{\partial x^j} \right)_p^* . \]
I.e. the dual basis to the coordinate vectors is \( \{ dx^1_p, \ldots, dx^n_p \} \). Thus, we can locally express any covariant vector \( \omega_p \in T^*_p M \) as
\[ \omega_p = \sum_{j=1}^n \omega^j_p dx^j_p. \]

Using this language, we can restate Proposition 6.5 as follows: the transformation laws for contravariant and covariant vector fields are mediated by the following formulas for changing coordinate vectors:
\[
\frac{\partial}{\partial x^j} \bigg|_p = \sum_{i=1}^n \frac{\partial y^i}{\partial x^j}(\hat{p}) \frac{\partial}{\partial y^i} \bigg|_p, \quad dx^j_p = \sum_{i=1}^n \frac{\partial x^j}{\partial y^i}(\hat{p}) dy^i_p. \tag{6.7}
\]

6.3. The Cotangent Bundle, and Covariant Vector Fields. We now want to glue together covariant vectors \( \omega_p \) at different points into a continuous object. The procedure will mirror what we did in Sections 3.5, 4.1, and 4.3 quite closely, and so we will be a lot more brief with the exposition here.

**Definition 6.7.** Let \( M \) be a smooth manifold. Its cotangent bundle \( T^* M \) is the disjoint union of all cotangent spaces
\[ T^* M = \bigcup_{p \in M} T^*_p M. \]
There is a natural projections \( \pi: T^* M \to M \) given by \( \pi(\omega_p) = p \) for any \( \omega_p \in T^*_p M \).

As with the tangent bundle, we can imbue \( T^* M \) with the structure of a smooth \( 2n \)-dimensional manifold in essentially the same way: given a chart \( (U, \varphi) \) for \( M \), we define a chart \( (\tilde{U}, \tilde{\varphi}) \) for \( T^* M \) as follows: \( \tilde{U} = \pi^{-1}(U) \), and if the coordinate functions of \( \varphi \) are \( (x^1, \ldots, x^n) \), then we define \( \tilde{\varphi} \) on any covariant vector \( \omega_p = \sum_{j=1}^n \omega^j dx^j_p \) by
\[ \tilde{\varphi} \left( \sum_{j=1}^n \omega^j dx^j_p \right) = (x^1(p), \ldots, x^n(p), \omega^1_p, \ldots, \omega^n_p) = (\varphi(p), \omega^1_p, \ldots, \omega^n_p). \]

The proof that these charts define a smooth structure on \( T^* M \) is almost identical to the proof of Proposition 3.25 (the corresponding statement for the tangent bundle), simply using the transformation law (6.7) for covariant vector fields (instead of contravariant vector fields) to show that the transition maps are smooth. We leave the details to the reader.

**Remark 6.8.** As differentiable manifolds, \( TM \) and \( T^* M \) are “the same”: they are diffeomorphic. We do not have the tools needed yet to prove this, but the idea is just a global version of the proof that \( V \cong V^* \). As in that case, there isn’t a “natural” diffeomorphism, unless extra structure is present. Following Remark 6.1 if we endow \( V \) with a fixed inner product, then there is a natural isomorphism \( V \leftrightarrow V^* \) given by \( v \leftrightarrow \langle \cdot, v \rangle \). Similarly, suppose that, for each \( p \), there is an inner product \( g_p: T_p M \times T_p M \to \mathbb{R} \), which varies smoothly in \( p \) (meaning that the map \( p \mapsto g_p(X_p, Y_p) \) is smooth for any smooth vector fields \( X, Y \in \mathcal{X}(M) \)). Such a function \( g: \mathcal{X}(M) \times \mathcal{X}(M) \to \mathbb{R} \) is called a **Riemannian metric**. If such a \( g \) exist (which is true on any smooth manifold, though we won’t prove that presently), then it yields a map \( TM \to T^* M \) defined by \( X_p \mapsto g_p(\cdot, X_p) \) which
is easily seen to be a bijection, and the smoothness of \( g \) makes it into a diffeomorphism. In fact, it is more: it is a generalization of what we called a global trivialization, or a \textit{bundle isomorphism}. The map commutes with the projections, and its restriction to each fibre is a linear isomorphism.

It is important to note that, although \( TM \) and \( T^*M \) are diffeomorphic, even in a strong bundle sense, they are still different objects. And, as with \( V \) and \( V^* \), they are not \textit{naturally} isomorphic. As we will shortly see, the cotangent bundle is fundamentally the better object to work with.

A \textit{section} of \( T^*M \) is a continuous function \( \omega : M \to T^*M \) with the property that \( \pi \circ \omega = \text{Id}_M \): i.e. it is a continuous choice of a covariant vector at each point. We call such sections (rough) \textit{contravariant vector fields}. As with contravariant vector fields, covariant vector fields are a module over \( C^\infty(M) \) (in fact over \( \text{Fun}(M, \mathbb{R}) \), where \((f \omega)_p(X_p) = f(p)\omega_p(X_p)\) and \((\omega + \lambda)_p = \omega_p + \lambda_p\).

Since \( T^*M \) is a smooth manifold, we can insist that a section be smooth.

\textbf{Definition 6.9.} The set of smooth covariant vector fields on \( M \) is denoted by \( \mathcal{X}^\ast(M) \), or alternatively by \( \Omega^1(M) \). The second notation goes along with the more common name for such vector fields: \textit{differential 1-forms}, or simply \textit{1-forms}.

Given a chart \((U, \varphi = (x^i)_{j=1}^n) \) in \( M \), the \textit{coordinate 1-forms} are \( dx^1, \ldots, dx^n \) (i.e. the sections \( p \mapsto dx^1_p \)); it is immediate from the definition of the smooth structure of \( T^*M \) that these are smooth.

In general, any (rough) covariant vector field \( \omega \in \mathcal{X}^\ast(M) \) can be expressed locally as

\[
\omega|_U = \sum_{j=1}^{n} \omega_j \, dx^j
\]

where \( \omega_j : M \to \mathbb{R} \) are the functions \( \omega_j(p) = \omega_p(\frac{\partial}{\partial x^j}) \).

Given a (rough) covariant vector field \( \omega \), we can use it to eat a (rough) vector field \( X \) to get a function on \( M \):

\[
\omega(X) : M \to \mathbb{R}, \quad \omega(X)(p) = \omega_p(X_p).
\]

As before, this allows us to think of covariant vector fields as certain kinds of functions. While a contravariant vector field \( X \in \mathcal{X}(M) \) can be identified with a derivation \( X : C^\infty(M) \to C^\infty(M) \), a covariant vector field \( \omega \in \mathcal{X}^\ast(M) \) can be identified with a linear map \( \omega : \mathcal{X}(M) \to \text{Fun}(M, \mathbb{R}) \), given by \((\omega(X))(p) = \omega_p(X_p)\) as above. In local coordinates, if we express \( \omega = \sum_{j=1}^{n} \omega_j \, dx^j \) and \( X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial x^i} \), then since \( dx^j_p = (\frac{\partial}{\partial x^j})_p^* \) we have \( \omega(X) = \sum_{j=1}^{n} \omega_j X^j \). This shows that, if the components of \( \omega \) and \( X \) are smooth, then so is the function \( \omega(X) \).

We can now characterize smoothness of a covariant vector field in terms of its coefficients, precisely mirroring Proposition 4.12 (the analogous statement for contravariant vector fields). The proof is very similar, and is left as a homework exercise.

\textbf{Proposition 6.10.} Let \( M \) be a smooth manifold, and let \( \omega : M \to T^*M \) be a (rough) covariant vector field. TFAE:

(a) \( \omega \) is smooth.
(b) The component functions of \( \omega \) in any chart are smooth.
(c) For every smooth contravariant vector field \( X \in \mathcal{X}(M) \), the function \( \omega(X) \) is smooth.
(d) Given any open subset \( U \subseteq M \) and \( X \in \mathcal{X}(U) \), the function \( \omega(X) : U \to \mathbb{R} \) is smooth.
The most important examples of smooth covariant vector fields are gradient fields, aka gradient 1-forms. Here we finally have the correct invariant generalization of the gradient from vector calculus.

**Definition 6.11.** Let \( f \in C^\infty(M) \). Following Definition 6.6 we have a covariant vector \( df_p \in T^*_p M \) for each \( p \). The **differential** of \( f \) \( df \) is the covariant vector field defined by \( df(p) = df_p \). Thinking of \( df \) as a linear function \( \mathcal{X}(M) \to \text{Fun}(M, \mathbb{R}) \), its action is then \( df(X) = X(f) \). If \( X \) is smooth, then \( X(f) \) is smooth, which means that \( df \in \Omega^1(M) \): it is a smooth covariant vector field.

Again, let us connect \( df \) with the earlier notation, where \( df : TM \to T\mathbb{R} \) is the map \( df(X_p) = df_p(X_p) \). That is: the differential of the map \( f : M \to \mathbb{R} \) is the linear map defined by \( (df(X_p))(g) = X_p(g \circ f) \). Now we think of \( df(X_p) \) as a real number rather than a vector in \( \mathbb{R} \) (which is just a real number), and the way to do this is to evaluate the derivation at \( \text{Id}_\mathbb{R} \). Once again, that connects the two notations here: we have

\[
\text{df}^{\text{new}}(X) = X(f) = X(\text{Id}_\mathbb{R} \circ f) = (\text{df}^{\text{old}}(X))(\text{Id}_\mathbb{R}).
\]

Referring to Definition 6.9 in local coordinates we can write

\[
df = \sum_{j=1}^n (df)_j dx^j, \quad \text{where} \quad (df)_j(p) = df_p \left( \frac{\partial}{\partial x^j} \right)_p = \frac{\partial f}{\partial x^j}(\hat{p}),
\]

meaning that

\[
df = \sum_{j=1}^n \frac{\partial f}{\partial x^j} dx^j.
\]

Here are some elementary properties of gradient 1-forms, whose proofs are left as exercises.

**Proposition 6.12.** Let \( M \) be a smooth manifold, and let \( f, g \in C^\infty(M) \).

(a) For constants \( a, b \in \mathbb{R} \), \( df(a f + b g) = a df + b dg \).

(b) \( df(fg) = f df g + g df f \).

(c) If \( \text{Im}(f) \subset (a, b) \) and \( h : (a, b) \to \mathbb{R} \) is smooth, then \( dh \circ f = (h' \circ f) df \).

(d) If \( f \) is constant, then \( df = 0 \).

The converse of item (d) above is true, and it one important application of gradient 1-forms.

**Proposition 6.13.** Suppose \( f \in C^\infty(M) \). Then \( df = 0 \) if and only if \( f \) is locally constant: it is constant on each component of \( M \).

**Proof.** We will assume \( f \) is connected; then the statement is \( df = 0 \) iff \( f \) is constant. Proposition 6.12(d) shows the “if” part of this equivalence, so we now treat the “only if” part. Suppose \( df = 0 \). Fix a point \( p \in M \), and let \( \mathcal{C} = f^{-1}(f(p)) = \{ q \in M : f(q) = f(p) \} \), which is closed by the continuity of \( f \). For any point \( q \in \mathcal{C} \), let \( (U, \varphi = (x^j)_{j=1}^n) \) be a chart at \( q \), and let \( r \in U \). In this chart, \( df = \sum_{j=1}^n \frac{\partial f}{\partial x^j} dx^j \). As \( \{dx^j\}_{j=1}^n \) form a basis of \( T^*_r M \), in order for \( df_r = 0 \), it therefore follows that \( \frac{\partial f}{\partial x^j}(\hat{r}) = 0 \) for \( 1 \leq j \leq n \). Thus, \( \frac{\partial f}{\partial x^j} \equiv 0 \) on \( U \) for all \( j \); by elementary calculus, this means \( f \) is constant on \( U \). This shows that \( \mathcal{C} \) is open. Since \( M \) is connected, it follows that \( \mathcal{C} = M \), concluding the proof. \( \square \)
6.4. **Pullbacks of Covariant Vector Fields.** Let $M, N$ be smooth manifolds, and $F: M \to N$ a smooth map. Given a point $p \in M$, the (ordinary) differential map $dF_p$ is a linear transformation $T_pM \to T_{F(p)}N$. Thus, there is a dual (transpose) map

$$(dF_p)^*: T_{F(p)}^*N \to T_p^*M.$$ 

This is called the pullback by $F$ at $p$, or the cotangent map, or the codifferential. By the definition of the transpose, its action is simply

$$[(dF_p)^*(\omega_{F(p)}))(X_p)] = \omega_{F(p)}(dF_p(X_p)), \quad \omega_{F(p)} \in T_{F(p)}^*N, \ X_p \in T_pM.$$ 

That is: we pullback a covariant vector by letting it act on the push-forward of its argument. But it turns out that this notion is better, in the following sense. Recall that vector fields generally do not turn out that this notion is better, in the following sense. Recall that vector fields generally do not push-forward: although we can push forward any given vector, we cannot do it to a whole vector field consistently unless that map $F$ is a diffeomorphism. Not so with covariant vector fields.

**Definition 6.14.** Let $M, N$ be smooth manifolds, and let $F: M \to N$ be a smooth map. Let $\omega: N \to T^*N$ be a (rough) covariant vector field. The pullback of $\omega$ along $F$ is the (rough) covariant vector field $F^*\omega: M \to T^*M$ defined by

$$(F^*\omega)_p = dF_p^*(\omega_{F(p)}).$$

I.e. the action of $F^*\omega$ is

$$(F^*\omega)_p(X_p) = \omega_{F(p)}(dF_p(X_p)).$$

Note that pullback is a linear operation: $F^*(a\omega + b\lambda) = aF^*\omega + bF^*\lambda$ for $a, b \in \mathbb{R}$.

By pulling back a 1-form (which is already a dual object), we see there is no ambiguity about which point to evaluate the argument vector field at, and hence we can pull back globally even if $F$ is not a diffeomorphism. Of course, we should expect that if $\omega \in \mathcal{F}^*(M)$ is smooth then the pullback $F^*\omega$ is smooth. This is true. To prove it, the following computational lemma will be useful.

**Lemma 6.15.** Let $M, N$ be smooth manifolds and $F: M \to N$ a smooth map. Let $u: N \to \mathbb{R}$ be continuous, and let $\omega: N \to T^*N$ be a (rough) covariant vector field on $N$. Then

$$F^*(u\omega) = (u \circ F)F^*\omega.$$  \hspace{1cm} (6.8)

In addition, if $u \in C^\infty(M)$, then

$$F^*du = d(u \circ F).$$ \hspace{1cm} (6.9)

**Proof.** We compute from the definition: for any $p \in M$,

$$(F^*(u\omega))_p = dF_p^*((u\omega)_{F(p)}) = dF_p^*(u(F(p))\omega_{F(p)})$$

$$= u(F(p))dF_p^*(\omega_{F(p)})$$

$$= (u \circ F)(p)(F^*\omega)_p,$$

which proves (6.8). Similarly, for any $X_p \in T_pM$,

$$(F^*du)_p(X_p) = (dF_p^*(du_{F(p)}))(X_p) = du_{F(p)}(dF_p(X_p))$$

$$= dF_p(X_p)(u)$$

$$= X_p(u \circ F)$$

$$= d(u \circ F)_p(X_p)$$

which proves (6.9). 

\[ \square \]
**Proposition 6.16.** Let $M, N$ be smooth manifolds and $F: M \to N$ a smooth map. If $\omega: N \to T^*N$ is a continuous covariant vector field, then so is $F^*\omega$. Moreover, if $\omega \in \Omega^1(N) = \mathcal{X}^*(N)$ is a smooth 1-form, then so is $F^*\omega$.

**Proof.** Fix $p \in M$, and let $q = F(p)$. Let $(V, \psi = (y^j)_{j=1}^n)$ be a chart at $q$, and let $U = F^{-1}(V)$ (which is an open neighborhood of $p$). Then we may write $\omega|_U$ in local coordinates as $\sum_{j=1}^n \omega_j dy^j$. Applying (6.8), we then have

$$F^*\omega = F^* \left( \sum_{j=1}^n \omega_j dy^j \right) = \sum_{j=1}^n F^*(\omega_j dy^j) = \sum_{j=1}^n (\omega_j \circ F) F^*(dy^j).$$

Now, by (6.9), $F^*(dy^j) = d(y^j \circ F)$, which are smooth gradient forms. The coefficient functions $\omega_j \circ F$ are continuous if $\omega$ is continuous, and smooth of $\omega$ is smooth. This concludes the proof. □

**Example 6.17.** Let $F: \mathbb{R}^3 \to \mathbb{R}^2$ be the map $(u, v) = F(x, y, z) = (yz, e^{xy}z^2)$. Let $\omega \in \Omega^1(\mathbb{R}^2)$ be the 1-form $\omega = u dv + 2v du$. Then

$$F^*\omega = F^*(u dv) + 2F^*(v du) = (u \circ F)(v \circ F) + 2(v \circ F)(u \circ F)$$

$$= (yz)d(e^{xy}z^2) + 2(e^{xy}z^2)d(yz)$$

$$= yz[e^{xy}d(z^2) + z^2d(e^{xy})] + 2e^{xy}x^2[y dz + z dy]$$

$$= yze^{xy} \cdot 2z dx + z^2e^{xy}(x dy + y dx) + 2e^{xy}x^2(y dz + z dy)$$

$$= yz^2e^{xy} dx + (x^2 + 2x^2)z e^{xy} dy + 2(yz^2 + x^2y)e^{xy} dz.$$

**Example 6.18.** Let $M = \mathbb{R}^2_+$ be the right half-plane $\{(x, y) \in \mathbb{R}^2: x > 0\}$. Consider the smooth map $Id: M \to M$. Let us use Cartesian coordinates on $M$ in the codomain of $Id$ and polar coordinates on the domain. Taking $\omega = x dy - y dx$ on $M$, we then have

$$\omega = x dy - y dx = Id^*(x dy - y dx)$$

$$= (r \cos \theta)d(r \sin \theta) - (r \sin \theta)d(r \cos \theta)$$

$$= r \cos \theta \cdot (\sin \theta dr + r \cos \theta d\theta) - r \sin \theta \cdot (\cos \theta dr - r \sin \theta d\theta)$$

$$= (r \cos \theta \sin \theta - r \sin \theta \cos \theta) dr + (r^2 \cos^2 \theta + r^2 \sin^2 \theta) \cdot d\theta = r^2 d\theta.$$

Now, recall the definition of the Lie derivative of a vector field $Y$ with respect to $X$:

$$\mathcal{L}_X(Y) = \left. \frac{d}{dt} \right|_{t=0} (\theta_t)_*(Y), \quad \text{where } \theta \text{ is the flow of } X.$$

To be clear: this is a pointwise definition: $\mathcal{L}_X(Y)(p) = \left. \frac{d}{dt} \right|_{t=0} (\theta_t)_*(Y)(p)$, the latter being the derivative of a map from an interval in $\mathbb{R}$ into the vector space $T_p M$.

This suggests a way to define the Lie derivative of a 1-form with respect to a vector field: by pulling back along $\theta$ instead of pushing forward.

**Definition 6.19.** Let $\omega \in \Omega^1(M)$ be a 1-form, and let $X \in \mathcal{X}(M)$ be a smooth vector field. Define a new 1-form $\mathcal{L}_X \omega$ as follows:

$$\mathcal{L}_X \omega = \left. \frac{d}{dt} \right|_{t=0} (\theta_t)^* \omega$$

where $\theta$ is the flow of $X$. 
To show this is well-defined, we need to show that the limit exists, and indeed defines a (smooth) covariant vector field on \( M \). To be clear, the definition is
\[
\mathcal{L}_X \omega |_p = \left. \frac{d}{dt} \theta^*_t \omega |_p \right|_{t=0} = \lim_{t \to 0} \frac{d(\theta^*_t \omega |_{\theta_t(p)}) - \omega_p}{t}.
\]
Note that it makes sense to use \( t \) rather than \(-t\) since we are pulling back instead of pushing forward (either way, the difference is just a minus sign convention).

**Proposition 6.20.** For any vector field \( X \in \mathfrak{X}(M) \) and any 1-form \( \omega \in \Omega^1(M) \), \( \mathcal{L}_X \omega \) is a 1-form, and its action on vector fields is
\[
(\mathcal{L}_X \omega)(Y) = \mathcal{L}_X(\omega(Y)) - \omega(\mathcal{L}_X Y) = X(\omega(Y)) - \omega([X,Y]).
\]  

**Proof.** Similar to the calculations in Proposition 5.24. We will see a more general result later, called “Cartan’s magic formula”, then proves this as a special case. \( \square \)

Note, rearranging (6.10) yields
\[
\mathcal{L}_X (\omega(Y)) = (\mathcal{L}_X \omega)(Y) + \omega(\mathcal{L}_X Y).
\]
That is: \( \mathcal{L}_X \) is a derivation even with respect to the “product” \( \Omega^1(M) \times \mathfrak{X}(M) \to C^\infty(M) \) given by \( (\omega, Y) \mapsto \omega(Y) \).

**Corollary 6.21.** If \( f \in C^\infty(M) \) and \( X \in \mathfrak{X}(M) \), then
\[
\mathcal{L}_X (df) = d(\mathcal{L}_X f).
\]

**Proof.** This is just a calculation: fixing another vector field \( Y \), from Proposition 6.20 we have
\[
[L_X(df)](Y) = L_X(df(Y)) - df([X,Y]) = L_X(Y(f)) - [X,Y](f)
= XY(f) - (XY(f) - YX(f))
= YX(f) = [d(X(f))](Y).
\]
Since \( \mathcal{L}_X f = X(f) \), this concludes the proof. \( \square \)

Again, we will see later a very general result that Lie derivatives commute with “exterior” derivatives, of which \( d: f \mapsto df \) is a special case.

### 6.5. Line Integrals.
What are 1-forms? Primary, they are precisely those objects that can be integrated over curves (in a coordinate-independent fashion). Indeed, in the calculus integral
\[
\int_a^b f(t) \, dt
\]
we should really think of \( f(t) \, dt \) as a single object, a 1-form on \([a, b] \). (Since \([a, b] \) is not a manifold, what we really mean here is a 1-form on a slightly larger open interval restricted to \([a, b] \).) So, we take \( \omega = f(t) \, dt \) for some smooth \( f \) (which is a general 1-form), and define
\[
\int_{[a,b]} \omega \equiv \int_a^b f(t) \, dt.
\]
This is not just a new notation. For example, it gives a very nice form for the change of variables theorem.
Proposition 6.22. Let \( \omega \in \Omega^1([a, b]) \). Let \( \varphi : [c, d] \to [a, b] \) be a homeomorphism, whose restriction to \((c, d)\) is a diffeomorphism. Then \( \varphi \) is either strictly increasing or strictly decreasing. Moreover, we have

\[
\int_{[c, d]} \varphi^* \omega = \pm \int_{[a, b]} \omega,
\]

+ if \( \varphi \) is increasing, and − if \( \varphi \) is decreasing.

Proof. Denote the (global) coordinate on \([c, d]\) as \( s \). First, any continuous bijection from \([c, d]\) to \([a, b]\) is necessarily monotone; that is a diffeomorphism means \( \phi'(s) \neq 0 \) for any \( s \), and so it is strictly monotone. Now, we have

\[
(\varphi^* \omega)(s) = \varphi^*(f(t) \, dt) = f \circ \varphi(s) \, d(\varphi(s)) = f \circ \varphi(s) \varphi'(s) \, ds
\]

by Lemma 6.15. Now, if \( \varphi \) is increasing then \( \varphi(c) = a \) and \( \varphi(d) = b \), and so by the change of variables theorem from calculus

\[
\int_{[c, d]} \varphi^* \omega = \int_c^d f(\varphi(s))\varphi'(s) \, ds = \int_a^b f(t) \, dt = \int_{[a, b]} \omega.
\]

If, on the other hand, \( \varphi \) is decreasing, then \( \varphi(c) = b \) and \( \varphi(d) = a \), and we get instead \( \int_a^b f(t) \, dt = -\int_b^a f(t) \, dt \). \( \square \)

So, the change of variables theorem from calculus is really the statement

\[
\int_C \varphi^* \omega = \int_{\varphi(C)} \omega
\]

for sufficiently nice maps \( \varphi \). We will now see how this generalizes beyond intervals in \( \mathbb{R} \). In any smooth manifold, a piecewise smooth curve \( \alpha \) is a continuous map \( \alpha : [a, b] \to M \) for some nonempty interval \([a, b]\) with the property that there are finitely many points \( a = a_0 < a_1 < \cdots < a_k = b \) such that \( \alpha|_{[a_{j-1}, a_j]} \) is smooth for \( 1 \leq j \leq k \). Note, this is slightly stronger than insisting that \( \gamma \) be smooth on \((a_{j-1}, a_j)\); it means that \( \gamma \) has an extension to a smooth curve \( \tilde{\gamma}_j : (a_{j-1} - \epsilon, a_j + \epsilon) \to M \) for some \( \epsilon > 0 \) (although the actual value of \( \gamma \) may well disagree with this extension off the interval \([a_{j-1}, a_j]\)).

Now, let \( \alpha \) be a piecewise smooth curve in \( M \). If \( \alpha : [a, b] \to M \) is actually smooth, then we can pullback along \( \alpha : \alpha^* \omega \in \Omega^1([a, b]) \). This means that \( \alpha^* \omega = f(t) \, dt \) for some smooth \( f \), and so we know how to integrate it. We therefore define

\[
\int_\alpha \omega \equiv \int_{[a, b]} \alpha^* \omega.
\]

More generally, if \( \alpha \) is any piecewise smooth curve on \([a, b]\) that is smooth on the intervals \([a_{j-1}, a_j]\) for \( 1 \leq j \leq k \), then we define

\[
\int_\alpha \omega \equiv \sum_{j=1}^k \int_{[a_{j-1}, a_j]} \alpha^* \omega.
\]

This is called a line integral. Here are some basic properties that are elementary to prove from the definitions.

Proposition 6.23. Let \( M \) be a smooth manifold, and let \( \alpha : [a, b] \to M \) be a piecewise smooth curve. Let \( \omega, \omega_1, \omega_2 \in \Omega^1(M) \), and let \( c_1, c_2 \in \mathbb{R} \).
(a) For any \( c_1, c_2 \in \mathbb{R} \),
\[
\int_\alpha (c_1 \omega_1 + c_2 \omega_2) = c_1 \int_\alpha \omega_1 + c_2 \int_\alpha \omega_2.
\]

(b) If \( \alpha \) is a constant curve, then \( \int_\alpha \omega = 0 \).

(c) If \( a \leq c \leq b \) and \( \alpha_1 = \alpha\mid_{[a,c]} \) and \( \alpha_2 = \alpha\mid_{[c,b]} \) then
\[
\int_\alpha \omega = \int_{\alpha_1} \omega + \int_{\alpha_2} \omega.
\]

(d) Let \( N \) be a smooth manifold and \( F : M \to N \) a smooth map. For any \( \eta \in \Omega^1(N) \),
\[
\int_\alpha F^*\eta = \int_{F \circ \alpha} \eta.
\]

Proof. We prove (d), leaving the others as calculations to the reader. First, it suffices to prove in
the case that \( \alpha \) is smooth, since in general if \( \alpha \) is smooth on the intervals \([a_j-1, a_j]\) then \( F \circ \alpha \) is smooth on the same intervals. Thus, if \( \alpha : [a, b] \to M \) is smooth, we have
\[
\int_\alpha F^*\eta = \int_a^b \alpha^*F^*\eta.
\]

Note that, for any \( X \in \mathfrak{X}([a, b]) \), then for any \( t \in [a, b] \)
\[
(\alpha^*F^*\eta)(X)(t) = (F^*\eta)(d\alpha_t(X_t))(\alpha(t)) = \eta_{F(\alpha(t))}(dF_{\alpha(t)} \circ d\alpha_t(X_t))
= \eta_{F(\alpha(t))}(d(F \circ \alpha)_t(X_t))
= (F \circ \alpha)^*\eta(X)(t).
\]
I.e. \( \alpha^*F^*\eta = (F \circ \alpha)^*\eta \), and so
\[
\int_\alpha F^*\eta = \int_a^b (F \circ \alpha)^*\eta = \int_{F \circ \alpha} \eta.
\]

Example 6.24. Let \( M = \mathbb{R}^2 \setminus \{0\} \), and define \( \omega \in \Omega^1(M) \) in local (global) coordinates by
\[
\omega = \frac{x \, dy - y \, dx}{x^2 + y^2}.
\]
Let \( \alpha : [0, 2\pi] \to M \) be the standard parametrization of the ccw unit circle \( \alpha(t) = (\cos t, \sin t) \).
Then we compute
\[
\alpha^*\omega = \frac{\cos t \sin t - \sin t \cos t}{\cos^2 t + \sin^2 t} = \cos t \cos t \, dt - \sin t(- \sin t) \, dt = dt.
\]
Thus
\[
\int_\alpha \omega = \int_0^{2\pi} dt = 2\pi.
\]

We have defined curves to have a parametrization as part of the definition. But it turns out
that the dependence on the parametrization is very mild. A curve \( \beta : [c, d] \to M \) is called a
reparametrization of \( \alpha : [a, b] \to M \) if there is a homeomorphism \( \varphi : [c, d] \to [a, b] \) so that
\( \beta = \alpha \circ \varphi \), and moreover if \( \alpha \) is smooth on \([a_j-1, a_j]\) then \( \varphi^{-1}\mid_{[a_j-1, a_j]} \) is a diffeomorphism. Such
a map is therefore monotone everywhere, and strictly monotone on each interval \( \varphi[a_j-1, a_j] \). We
call this a **forward reparametrization** if \( \varphi \) is increasing, and a **backward reparametrization** if \( \varphi \) is decreasing.

**Proposition 6.25.** Let \( M \) be a smooth manifold and \( \omega \in \Omega^1(M) \). Let \( \alpha \) be a piecewise smooth curve, and let \( \beta \) be a reparametrization. Then

\[
\int_\beta \omega = \pm \int_\alpha \omega,
\]

+ if it is a forward reparametrization and – if it is backward.

**Proof.** Again, it suffices to assume \( \alpha \) is smooth (by restricting to its smooth subintervals and adding up on both sides at the end). Then we have

\[
\int_\beta \omega = \int_{[c,d]} (\alpha \circ \varphi)^* \omega = \int_{[c,d]} \varphi^*(\alpha^* \omega) = \pm \int_{[a,b]} \alpha^* \omega = \pm \int_\alpha \omega
\]

by Proposition 6.22. \( \square \)

Thus, the line integral does not really depend on the curve \( \alpha \) by only its image \( \alpha[a,b] \), together with an “orientation” (forward or backward).

Let us now connect our pullback definition of line integrals with the usual definition given in vector calculus.

**Proposition 6.26.** Let \( \alpha : [a, b] \to M \) be a piecewise smooth curve, and let \( \omega \in \Omega^1(M) \). Then

\[
\int_\alpha \omega = \int_a^b \omega_{\alpha(t)}(\dot{\alpha}(t)) \, dt.
\]

**Proof.** First, note that \( \alpha[a, b] \subset M \) is compact, and so there are finitely many charts \( (U_j, \varphi_j) \) that cover the image. Subdividing the domain further if necessary, we may assume that \( \alpha \) is smooth on the intervals \( [a_j-1, a_j] \) where \( \alpha[a_{j-1}, a_j] \subset U_j \). It therefore suffices to assume that \( \alpha \) is smooth, and its image is contained in a single chart \( (U, \varphi) \). Let \( \dot{\alpha}(t) = \varphi \circ \alpha(t) = (\alpha^1(t), \ldots, \alpha^n(t)) \), and write \( \omega = \sum_{j=1}^n \omega_j \, dx^j \) in this chart. Then we have

\[
\omega_{\alpha(t)}(\dot{\alpha}(t)) = \sum_{j=1}^n \omega_j(\alpha(t)) \, dx^j(\dot{\alpha}(t)) = \sum_{j=1}^n \omega_j(\alpha(t)) \, \dot{\alpha}^j(t).
\]

Now \( \alpha^j : [a, b] \to \mathbb{R} \) is a calculus function, and so \( d\alpha^j_t = \dot{\alpha}_j(t) \, dt \). Thus we have

\[
\omega_{\alpha(t)}(\dot{\alpha}(t)) \, dt = \sum_{j=1}^n (\omega_j \circ \alpha)(t) d(\alpha^j_t).
\]

On the other hand

\[
(\alpha^* \omega)_t = \sum_{j=1}^n \alpha^*(\omega_j dx^j) = \sum_{j=1}^n (\omega_j \circ \alpha)(t) d(x^j \circ \alpha)(t) = \sum_{j=1}^n (\omega_j \circ \alpha)(t) d(\alpha^j)_t
\]

by Lemma 6.15. Thus

\[
\int_\alpha \omega = \int_{[a,b]} \alpha^* \omega = \int_{[a,b]} \omega_{\alpha(t)}(\dot{\alpha}(t)) \, dt.
\]

\( \square \)

This brings us to the Fundamental Theorem of Calculus.
**Theorem 6.27.** Let $M$ be a smooth manifold. Let $f \in C^\infty(M)$ and let $\alpha: [a, b] \to M$ be a piecewise smooth curve. Then

$$\int_\alpha df = f(\alpha(b)) - f(\alpha(a)).$$

**Proof.** Let $a = a_0 < a_1 < \cdots < a_k = b$ be partitions points so that $\alpha$ is smooth on $[a_{j-1}, a_j]$ for $1 \leq j \leq k$. Let $\alpha_j = \alpha|_{[a_{j-1}, a_j]}$. By Proposition 6.26, we have

$$\int_{\alpha_j} df = \int_{a_{j-1}}^{a_j} df(\alpha(t)) \, dt.$$

Now we compute

$$df(\alpha(t))(\dot{\alpha}(t)) = \left. \frac{d}{ds} \right|_{s=t} (f \circ \alpha) = (f \circ \alpha)'(t).$$

Thus

$$\int_{\alpha_j} df = \int_{a_{j-1}}^{a_j} (f \circ \alpha)'(t) \, dt.$$

The function $t \mapsto f \circ \alpha$ is smooth, and hence we may apply the classical fundamental theorem of calculus to compute this is equal to

$$\int_{\alpha_j} df = f(\alpha(a_j)) - f(\alpha(a_{j-1})).$$

Finally, we then have

$$\int_\alpha df = \sum_{j=1}^k \int_{\alpha_j} df = \sum_{j=1}^n [f(\alpha(a_j)) - f(\alpha(a_{j-1}))] = f(\alpha(a_k)) - f(\alpha(a_0))$$

as this is a telescoping sum, concluding the proof.

---

### 6.6. Exact and Closed 1-Forms

In classic vector calculus, a vector field is called **conservative** if it is the gradient of a smooth function. We might use the same word for 1-forms, but the more common term is **exact**: a 1-form $\omega \in \Omega^1(M)$ is exact if there is some $f \in C^\infty(M)$ for which $\omega = df$. By the Fundamental Theorem of Calculus (Theorem 6.27), if $\omega$ is exact then $\int_\alpha \omega$ only depends on $\alpha$ through its endpoints: i.e. if $\alpha$ and $\beta$ are any two piecewise smooth curves connecting $p$ to $q$, then $\int_\alpha \omega = \int_\beta \omega$. In particular, for any closed curve $\alpha$, $\int_\alpha \omega = 0$. In fact, this is an equivalence. To see this, we first need a path-connectedness lemma.

**Lemma 6.28.** If $M$ is a connected smooth manifold, and $p, q \in M$, there is a piecewise smooth curve $\alpha: [a, b] \to M$ with $\alpha(a) = p$ and $\alpha(b) = \beta$.

We could always choose $\{a, b\} = \{0, 1\}$, but it is convenient to have the freedom to use other parameter domains.

**Proof.** This is a (by now) standard local-to-global-connectedness proof. Fix $p \in M$, and let $C \subseteq M$ be the set of points connected to $p$ via some piecewise smooth curve:

$$C = \{q \in M : \exists a < b \text{, piecewise smooth } \alpha: [a, b] \to M \text{ s.t. } \alpha(a) = p, \alpha(b) = q\}.$$  

The constant curve $\alpha(t) = p$ shows that $p \in C$, so $C \neq \emptyset$. 

---
• \(\mathcal{C}\) is open: Fix some \(q \in \mathcal{C}\), and let \((U, \varphi)\) be a chart at \(q\). Fix \(\delta > 0\) so that \(B_\delta(\hat{q}) \subseteq \varphi(U)\), and let \(V = \varphi^{-1}(B_\delta(\hat{q}))\). For any \(r \in V\), the straight line curve \(\hat{\alpha}(t) = (1 - t)\hat{q} + t\hat{r}\) has image contained in \(B_\delta(\hat{q})\) (since this ball is convex), and has a smooth extension to a slightly longer line on both sides. Then \(\alpha = \varphi^{-1} \circ \hat{\alpha}\) is a smooth curve from \(q\) to \(r\); concatenating it with the curve connecting \(p\) to \(q\) (presumed to exists since \(q \in \mathcal{C}\)) shows that \(r \in \mathcal{C}\). Thus the neighborhood \(V \ni q\) is contained in \(\mathcal{C}\), showing that \(\mathcal{C}\) is open.

• \(\mathcal{C}\) is closed: let \(q \in \partial \mathcal{C}\). Again fix a chart \((U, \varphi)\) at \(q\). Since \(q \in \partial \mathcal{C}\), the open neighborhood \(U\) of \(q\) contains some point \(r \in \mathcal{C} \cap U\). Constructing the straight-line path in local coordinates just as above thus shows that \(r \in \mathcal{C}\). Thus \(\partial \mathcal{C} \subseteq \mathcal{C}\), so \(\mathcal{C}\) is closed.

Thus \(\mathcal{C}\) is a clopen nonempty subset of the connected manifold \(M\), and hence \(\mathcal{C} = M\). \(\square\)

**Theorem 6.29.** Let \(M\) be a smooth manifold, and let \(\omega \in \Omega^1(M)\). Then \(\omega\) is exact iff \(\int_\alpha \omega = 0\) for any closed piecewise smooth curve \(\alpha\).

**Proof.** Using the discussion above, the Fundamental Theorem of Calculus proves the ‘only if’ direction of the theorem. For the converse, suppose integrals of \(\omega\) around closed curves are always 0. This in fact implies that such integrals are path independent: for if \(\alpha\) and \(\beta\) are two piecewise smooth curves connecting \(p\) to \(q\), then the reversed curve \(-\beta\) connects \(q\) to \(p\), and so the concatenation \(\alpha - \beta\) is a closed curve. Thus

\[
0 = \int_{\alpha - \beta} \omega = \int_\alpha \omega + \int_{-\beta} \omega = \int_\alpha \omega - \int_\beta \omega.
\]

Now, let us assume that \(M\) is connected. We may then define an “integral” operation for any two points \(p, q \in M\):

\[
\int_p^q \omega = \int_\alpha \omega \quad \text{for any piecewise smooth curve } \alpha \text{ connecting } p \text{ to } q.
\]

This is well-defined and makes sense for any \(p, q\) by Lemma 6.28 Also, using Propositions 6.25 and 6.23(c), we have

\[
\int_p^q \omega = -\int_q^p \omega, \quad \int_p^r \omega = \int_p^q \omega + \int_q^r \omega.
\]

So: fix a base point \(p_0 \in M\), and define a function \(f : M \to \mathbb{R}\) by \(f(p) = \int_{p_0}^p \omega\). We will show that \(f \in C^\infty(M)\), and \(df = \omega\). To accomplish this, let \(q \in M\), and fix a chart \((U, \varphi = (x^j)_{j=1}^n)\) at \(q\).

In this chart we have \(\omega = \sum_{j=1}^n \omega_j \, dx^j\). We will compute that \(\omega_j(q) = \frac{\partial f}{\partial x^j}(q)\), which means that \(df_q = \omega_q\) as required.

Fix \(j \in \{1, \ldots, n\}\) and let \(\epsilon > 0\) be small enough that the curve \(\alpha : (-\epsilon, \epsilon) \to U\) given by \(\varphi \circ \alpha(t) = (0, \ldots, t, \ldots, 0)\) (with the \(t\) in the \(j\)th place) stays contained in \(U\). Let \(p_1 = \alpha(-\epsilon)\), and define a new function \(\tilde{f}(p) = \int_{p_1}^p \omega\). Then

\[
\tilde{f}(p) - f(p) = \int_{p_0}^p \omega - \int_{p_1}^p \omega = \int_{p_0}^{p_1} \omega + \int_{p}^{p_1} \omega = \int_{p_0}^{p_1} \omega
\]

which is a constant. Hence, it suffices to show that \(\tilde{f}\) is smooth and satisfies \(\frac{\partial \tilde{f}}{\partial x^j}(q) = \omega_j(q)\). Well, by construction \(\hat{\alpha}(t) = \frac{\partial}{\partial x^j} \circ \alpha(t)\), and so

\[
\omega_{\alpha(t)}(\hat{\alpha}(t)) = \sum_{i=1}^n \omega_i(\alpha(t)) \, dx^i \left( \frac{\partial}{\partial x^j} \bigg|_{\alpha(t)} \right) = \omega_j(\alpha(t)).
\]
Hence, we have
\[ f(\alpha(t)) = \int_{p_1}^{\alpha(t)} \omega = \int_{-\epsilon}^{t} \omega_{\alpha(s)}(\dot{\alpha}(s)) \, ds = \int_{-\epsilon}^{t} \omega_j(\alpha(s)) \, ds \]

But \( \omega \in \Omega^1(M) \), so its components are smooth. It then follows from the (classical) fundamental theorem of calculus that \( \tilde{f} \circ \alpha \) is smooth, and
\[ \frac{\partial \tilde{f}}{\partial x^j}(q) = \dot{\alpha}(0) \tilde{f} = \left. \frac{d}{dt} \tilde{f} \circ \alpha(t) \right|_{t=0} = \frac{d}{dt} \int_{-\epsilon}^{t} \omega_j(\alpha(s)) \, ds = \omega_j(\alpha(0)) = \omega_j(q). \]

Since the components \( \omega_j \) are smooth functions of \( q \), this shows that \( \tilde{f} \) is smooth, and moreover that \( df_q = df_{\tilde{q}} = \omega \) as claimed.

Finally, if \( M \) is not connected, then the above argument shows there are functions \( f_i \in C^\infty(M_i) \) on the connected components \( M_i \) of \( M \) so that \( df_i = \omega \big|_{M_i} \). Then the function \( f \) whose value on \( M_i \) is \( f_i \) is smooth and satisfies \( df = \omega \).

Now, not every 1-form is exact: as Example [6.24] shows, there are 1-forms with non-zero integrals around closed curves. So how can we tell if a 1-form \( \omega \) is exact, without already knowing a function \( f \) for which \( \omega = df \)? To see the answer, we work in local coordinates: write \( \omega = \sum_{j=1}^{n} \omega_j \, dx^j \) in some chart. If \( \omega = df \), this means that \( \omega_j = \partial f/\partial x^j \). Since \( f \) is smooth, its mixed partials commute, which means that
\[ \frac{\partial \omega_j}{\partial x^k} = \frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i} = \frac{\partial \omega_i}{\partial x^j}. \]

**Definition 6.30.** A 1-form \( \omega \) is called closed if, in every chart, the components \( \omega_j \) satisfy
\[ \frac{\partial \omega_j}{\partial x^i} = \frac{\partial \omega_i}{\partial x^j}. \quad (6.11) \]

The preceding calculation shows that every exact 1-form is closed. This gives us a negative test for exactness: if we can find some chart in which \( (6.11) \) fails at some point, then \( \omega \) is not exact. This is still a tall order, however: it, in principle, requires us to check every chart, which is impractical. In fact, as we will show next, it suffices to check this condition in any one chart; moreover, there is an equivalent invariant condition.

**Proposition 6.31.** For any \( \omega \in \Omega^1(M) \), TFAE:

(a) \( \omega \) is closed.
(b) \( \omega \) satisfies \( (6.11) \) in some chart at each point.
(c) Given any open \( U \subseteq M \) and any vector fields \( X, Y \in \mathfrak{X}(U) \), we have
\[ X(\omega(Y)) - Y(\omega(X)) = \omega([X,Y]). \quad (6.12) \]

**Proof.** The implication (a) \( \implies \) (b) is immediate from the definition (which states that \( (6.11) \) holds in every chart). Now, assume (b) holds true, and fix any open set \( U \) and vector fields \( X, Y \in \mathfrak{X}(U) \). Fix a point \( p \in U \), and let \( (V, \varphi = (x^i)_1^n) \) be a coordinate chart at \( p \), contained in \( U \), in which \( (6.11) \) holds. Expand \( \omega = \sum_{i=1}^{n} \omega_i \, dx^i \), \( X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial x^i} \), and \( Y = \sum_{k=1}^{n} Y^k \frac{\partial}{\partial x^k} \), and compute
\[ X(\omega(Y)) = X \left( \sum_{i=1}^{n} \omega_i Y^i \right) = \sum_{i=1}^{n} (\omega_i X(Y^i) + Y^i X(\omega_i)) = \sum_{i=1}^{n} \left( \omega_i X(Y^i) + \sum_{j=1}^{n} Y^j X_j \frac{\partial \omega_i}{\partial x^j} \right). \]
Exchanging the roles of $X$ and $Y$ shows that

$$Y(\omega(X)) = \sum_{i=1}^{n} \left( \omega_i Y(X^i) + \sum_{j=1}^{n} X^i Y_j \frac{\partial \omega_i}{\partial x^j} \right).$$

Subtracting, this gives

$$X(\omega(Y)) - Y(\omega(X)) = \sum_{i=1}^{n} \omega_i [X(Y^i) - Y(X^i)] + \sum_{i,j=1}^{n} Y^i X^j \left( \frac{\partial \omega_i}{\partial x^j} - \frac{\partial \omega_j}{\partial x^i} \right),$$

and the last term is 0 by assumption (of (6.11)). We also note that

$$[X, Y] = XY - YX = X \left( \sum_{i=1}^{n} Y^i \frac{\partial}{\partial x^i} \right) - Y \left( \sum_{i=1}^{n} X^i \frac{\partial}{\partial x^i} \right) = \sum_{i=1}^{n} [X(Y^i) - Y(X^i)] \frac{\partial}{\partial x^i},$$

and so $\sum_{i=1}^{n} \omega_i [X(Y^i) - Y(X^i)] = \omega([X, Y])$, concluding the calculation, and verifying that (b) $\implies$ (c).

Finally, we proved (c) $\implies$ (d). Fix any chart, and expand the 1-form $\omega$ and the vector fields $X, Y$ in coordinate bases. Note that the calculation of (6.13) and the following equation show that, in general,

$$X(\omega(Y)) - Y(\omega(X)) = \omega([X, Y]) + \sum_{i,j=1}^{n} Y^i X^j \left( \frac{\partial \omega_i}{\partial x^j} - \frac{\partial \omega_j}{\partial x^i} \right).$$

Thus, assuming (c), we have for any smooth functions $X^i$ and $Y^j$,

$$\sum_{i,j=1}^{n} Y^i X^j \left( \frac{\partial \omega_i}{\partial x^j} - \frac{\partial \omega_j}{\partial x^i} \right) = 0.$$

Choosing the functions so that $Y^i = X^j = 1$ and $Y^k = 0$ if $k \neq i$ and $X^k = 0$ if $k \neq j$ yields the result.

\[\square\]

**Remark 6.32.** We will soon talk about $k$-forms for $k > 1$, and extend the operator $d$ to act on all forms. In particular, for a 1-form $\omega$, $d\omega$ will be a 2-form (which eats two vector fields); its action will be defined as

$$d\omega(X, Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y]).$$

So, the condition that $\omega$ is closed is really the condition that $d\omega = 0$; hence, we have if $\omega = df$ then $d\omega = 0$, which is to say that $d^2 = 0$. This generalizes the classical vector calculus statement that $\nabla \times \nabla f = 0$ for any smooth function (the curl of a gradient is 0).

So, we have a nice, easy to check, invariant condition that gives a half-decidable test for whether a given 1-form is exact. It is not, however, a fully-decidable test.

**Example 6.33.** Consider again the 1-form of Example 6.24:

$$\omega = -\frac{y}{x^2 + y^2} \, dx + \frac{x}{x^2 + y^2} \, dy.$$  

We showed (using the fundamental theorem of calculus) that $\omega$ is not exact. Nevertheless, we can quickly calculate that

$$\frac{\partial}{\partial y} \left( -\frac{y}{x^2 + y^2} \right) = \frac{-x^2 + y^2}{(x^2 + y^2)^2} = \frac{\partial}{\partial x} \left( \frac{x}{x^2 + y^2} \right).$$

Thus $\omega$ is closed, but not exact. More on this shortly.
Both closeness and exactness are natural conditions, in the following sense.

**Corollary 6.34.** Let \( F: M \rightarrow N \) be a smooth map between manifolds, and let \( \omega \in \Omega^1(N) \). If \( \omega \) is exact, then \( F^*\omega \) is exact. Moreover, if \( F \) is a diffeomorphism, and \( \omega \) is closed, then \( F^*\omega \) is closed.

**Proof.** First, by Lemma 6.15, if \( \omega \) is exact and so has the form \( \omega = df \), then \( F^*(\omega) = F^*(df) = d(f \circ F) \) is also exact. Now, suppose we only know \( \omega \) is closed. One can compute that, for any vector fields \( X, Y \in \mathfrak{X}(M) \) and \( \omega \in \Omega^1(N) \),

\[
X(F^*\omega(Y)) = F_*X(\omega(F_*Y)).
\]

Thus, we simply have

\[
X(F^*\omega(Y)) - Y(F^*\omega(X)) = F_*X(\omega(F_*Y)) - F_*Y(\omega(F_*X)).
\]

Since \( \omega \) is closed on \( N \), it follows that

\[
F_*X(\omega(F_*Y)) - F_*Y(\omega(F_*X)) = \omega([F_*X, F_*Y]) = \omega(F_*[X, Y])
\]

and one final calculation shows that this equals \( F^*\omega([X, Y]) \), proving that \( F^*\omega \) is closed.

Alternatively, it is actually simpler to work in local coordinates here. Let \((U, \varphi)\) be a chart for \( N \). Then \((F^{-1}(U), \varphi \circ F)\) is a chart for \( M \). In these coordinates, the map \( F \) is given by the identity, and so it is clear that \( \omega \) satisfies \((6.11)\) in the \( \varphi \) coordinates iff \( F^*\omega \) satisfies \((6.11)\) in the \( \varphi \circ F \)-coordinates. The result then follows by Proposition 6.31. \( \square \)

**Remark 6.35.** (1) In fact, in the local proof of Corollary 6.34, one only needs that \( F \) is a *local* diffeomorphism: in this case, an invariant form of the inverse function theorem will allow the local-coordinates argument to work on some sufficiently small open subset of the original chart \( U \). We have yet to state and prove such an invariant inverse function theorem, so we leave it as is for now.

(2) One might wonder whether the local diffeomorphism condition is really necessary to preserved closedness, and in fact it is *not* necessary: like exactness, closedness is invariant under all pullbacks. The reason is the same: referring to Remark 6.32, \( \omega \) is closed iff \( d\omega = 0 \), and then we will have \( dF^*\omega = F^*d\omega = F^*(0) = 0 \), as with functions. In the next section, we will develop the technology to prove this.

Now, the question of how far apart the conditions *exact* and *closed* are is a deep and interesting one. Returning once more to Example 6.24, although this form is not exact on \( \mathbb{R}^2 \setminus \{0\} \), if we restrict it to \( M = \mathbb{R}^2 \setminus \{(x, 0): x \in \mathbb{R}\} \), then the function \( M = \tan^{-1} \frac{y}{x} \) is smooth on \( M \), and \( \omega = df \). (In fact, although the formula doesn’t make sense, this function \( f \) has a smooth extension to any domain \( \mathbb{R}^2 \setminus \gamma \) for any ray \( \gamma = \{tv: t > 0\} \) where \( v \) is some nonzero vector in \( \mathbb{R}^2 \).) What this demonstrates is that exactness is not really a local condition (while closedness is): it depends on the global topology of the manifold where the form is defined. The main basic result in this direction is Poincaré’s lemma, which shows that closed 1-forms on star-shaped regions in \( \mathbb{R}^n \) are indeed exact.

**Theorem 6.36 (Poincaré Lemma).** Let \( U \subseteq \mathbb{R}^n \) be open and star-shaped. Then every closed 1-form on \( U \) is exact.

**Note:** star-shaped means there is a “center” point \( c \in U \) so that, for every \( p \in U \), the line segment \( \{(1 - t)c + tp: 0 \leq t \leq 1\} \) is contained in \( U \). For example, convex sets are star-shaped.

**Proof.** Let \( \omega \in \Omega^1(U) \) be closed. First, let \( T: \mathbb{R}^n \rightarrow \mathbb{R}^n \) be the translation \( T(x) = x - c \); since \( T \) is a diffeomorphism, \( T^* \) preserves closedness and exactness, and so it suffices to prove that the
closed 1-form $T^*\omega$ is exact. Henceforth we rename $T(U)$ as $U$ and $T^*\omega$ as $\omega$, and proceed under the assumption that $c = 0$ (which is just for notational convenience). Thus, for any $x \in U$, the straight-line $\alpha_x: [0, 1] \to U$ given by $\alpha_x(t) = tx$ is a smooth curve in $U$, and we can define a function $f: U \to \mathbb{R}$ by

$$f(x) = \int_{\alpha_x} \omega.$$ 

We will show that $f$ is smooth, and that $\omega = df$ i.e. that $\frac{\partial f}{\partial x^j} = \omega_j$ for $1 \leq j \leq n$.

Expand $\omega = \sum_{j=1}^n \omega_j \, dx^j$. Use Proposition 6.26 to write $f$ as

$$f(x) = \int_0^1 \omega_{\alpha_x(t)}(\dot{\alpha}_x(t)) \, dt = \sum_{j=1}^n \int_0^1 \omega_j(tx)x^j \, dt.$$ 

The integrand is a smooth function of $t, x^1, \ldots, x^n$, and so we may differentiate under the integral

$$\frac{\partial f}{\partial x^i}(x) = \sum_{j=1}^n \int_0^1 \left( t \frac{\partial \omega_j}{\partial x^i}(tx)x^j + \omega_j(tx)\delta^i_j \right) \, dt$$

$$= \int_0^1 \left( \sum_{j=1}^n \frac{\partial \omega_j}{\partial x^i}(tx)x^j + \omega_i(tx) \right) \, dt.$$ 

We now use the assumption that $\omega$ is closed, so that $\frac{\partial \omega_j}{\partial x^i} = \frac{\partial \omega_i}{\partial x^j}$, which gives

$$\frac{\partial f}{\partial x^i}(x) = \int_0^1 \left( \sum_{j=1}^n \frac{\partial \omega_i}{\partial x^j}(tx)x^j + \omega_i(tx) \right) \, dt = \int_0^1 \frac{d}{dt}(t\omega_i(tx)) \, dt = t\omega_i(tx)|_{t=0} = \omega_i(x).$$

Hence, the partial derivatives of $f$ are the (smooth) functions $\omega_i$. This shows $f$ is smooth, and that $df = \omega$, as claimed. \hfill \Box

**Remark 6.37.** The precise condition for this to work is much weaker than star-shaped: it is simply-connected. That is: it must be true that any two piecewise smooth curves from $p$ to $q$ are homotopic. If this is the case, then one can define $f$ as follows: choose a base point $p \in U$, and let $\alpha_x$ be some curve from $p$ to $x$ that is a finite collection of line segments in coordinate directions. Any two such curves are homotopic, and this (together with closedness of $\omega$) shows that $f$ is well-defined. A more involved version of the above calculation shows that $df(x) = \omega_x$.

**Corollary 6.38.** If $\omega \in \Omega^1(M)$ is closed, then every point $p \in M$ has a neighborhood $U$ so that $\omega|_U$ is exact.

**Proof.** Let $p \in M$, and choose a chart $(V, \varphi)$ at $p$; let $\hat{U} \subseteq \varphi V$ be a ball centered at $\varphi(p)$, and let $U = \varphi^{-1}(\hat{U})$. By assumption $\omega$ is closed and so, in the chart $(U, \varphi)$ condition (6.11) holds true. Thus, since $\hat{U}$ is convex, the Poincaré Lemma implies that the components $\omega_j$ of $\omega$ in the $\varphi$-coordinates $(x^j)_{j=1}^n$ satisfy $\omega_j = \frac{\partial f}{\partial x^j}$ for some smooth function $f$ on $\hat{U}$. It then follows (pulling back along $\varphi$) that $\omega|_U = df$ where $f = \hat{f} \circ \varphi$. Thus $\omega|_U$ is exact. \hfill \Box

**REFERENCES**
