Math 423 Course Notes

David Rose

October 18, 2003

Contents

1 Smooth Manifolds

1.1 Definitions and Examples

Before we give the definition of a manifold, we need a couple of preliminary definitions.

Definition 1.1. Chart
Let X be a topological space. An $\mathbb{R}^n$ **chart** on X is a map $\phi : U \to U'$ where
1) $U \subset X$ is open;
2) $U' \subset \mathbb{R}^n$ is open, and
3) ϕ is a homeomorphism.

Definition 1.2. Atlas
A C^∞ **atlas** in a topological space is a collection of charts $\{\phi_\alpha : U_\alpha \to U'_\alpha\}$ such that
1) $\{U_\alpha\}$ is an open cover of X, and
2) If $U_\alpha \cap U_\beta \neq \emptyset$, then $\phi_\beta \circ \phi_\alpha^{-1}(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$ is C^∞.

This last requirement says that change of coordinates should be smooth. In addition, we have the notion of equivalence: We say that two atlases are equivalent if their union is also an atlas. One can easily verify that this is indeed an equivalence relation.
Now, armed with these definitions, we are ready to define the notion of a smooth (C^∞) manifold.

Definition 1.3. Manifold
An **n-dimensional (C^∞) manifold** is a second countable Hausdorff topological space together with an equivalence class of C^∞ atlases into $\mathbb{R}^n$.

A few comments on this definition, which may be hard to digest at first: Foremost, manifold is a space that is "locally Euclidean" whose change of coordinates is smooth. It is important that change of coordinates are smooth, for, roughly speaking, if we want to do generalized calculus on manifolds, we will want to ensure that the calculus we do on one coordinate chart agrees with the calculus we do on any overlapping coordinate charts. Finally, Hausdorff and second countable ensure that the topology on our manifold is "nice" (for example, in a non-Hausdorff space, sequences can converge to two different limits!), and, as we will see later on, allows us to assert the existence of global objects when we can define them only locally. This will all be made more precise later on.
Now, let's look at some examples of manifolds.

Example 1.4.
Let $X = \mathbb{R}$, $U' = X = U$, and $\phi(x) = x$. This is the standard manifold structure on $\mathbb{R}$. $\square$

Example 1.5.
$\mathbb{C}^n$ is a manifold for all n. $\square$

Example 1.6.
If M is a manifold, and $V \subset M$ is open, then V is a manifold. $\square$

Example 1.7.
Note that $\mathrm{M_n}(\mathbb{R})$ is a manifold, for it can be identified with $\mathbb{R}^{n^2}$. Furthermore, the determinant map from $\mathrm{M_n}(\mathbb{R})$ to $\mathbb{R}$ is continuous, since it is a polynomial mapping from $\mathbb{R}^{n^2}$ to $\mathbb{R}$. Since $\mathrm{GL}(n, \mathbb{R}) = f^{-1}[(-\infty, 0) \cup (0, \infty)]$, it is an open set of $\mathrm{M_n}(\mathbb{R})$. So by the previous example, $\mathrm{GL}(n, \mathbb{R})$ is a manifold. $\square$

Example 1.8.
Let $X = S^2$, the two-sphere. Let's give X the subspace topology that it inherits as a subset of $\mathbb{R}^3$. First, we need to define charts. To do this, let $U_i^+ = \{x \in S^2 : x_i > 0\}$ and $U_i^- = \{x \in S^2 : x_i < 0\}$, which gives us an open cover of X, and define $\phi_1^{\pm}(x) = (x_2, x_3)$, $\phi_2^{\pm}(x) = (x_1, x_3)$, and $\phi_3^{\pm}(x) = (x_1, x_2)$.
Next, we need to verify that change of coordinates is smooth. Consider, for example, $\phi_2^+ \circ {\phi_1^+}^{-1}(u_1, u_2) = (\sqrt{1 - u_1^2 - u_2^2}, u_2)$, which is smooth in its region of definition. The other compositions yield similar results, so that it follows that X is indeed a manifold. $\square$

Example 1.9.
Now we consider a more non-trivial example of a manifold, real projective space. Specifically, let $\mathbb{R}P^{n-1}$ denote equivalence classes of lines in $\mathbb{R}^n$ through the origin, where two lines v and v' are equivalent if and only if there is a constant $\lambda \neq 0$ such that $v = \lambda v'$. Note that this is an equivalence relation.
The topology here is the quotient topology induced by the map $\pi : \mathbb{R}^n - \{0\} \to \mathbb{R}P^{n-1}$ defined by $(v) \mapsto [v]$. That is, $U \subset \mathbb{R}P^{n-1}$ is open if and only if $\pi^{-1}(U)$ is open in $\mathbb{R}^n - \{0\}$.
Charts here are given as follows:

$$U_i = \{[x_1, ..., x_n] \in \mathbb{R}P^{n-1} : x_i \neq 0\}$$

$$\phi_i : U_i \to \mathbb{R}^{n-1} \text{ is defined by } [x_1, ..., x_n] \mapsto \left(\frac{x_1}{x_i}, \cdots, \frac{x_n}{x_i}\right)$$

$$\phi_i^{-1} : (x_1, \cdots, x_{n-1}) \mapsto [x_1, \cdots, x_{i-1}, 1, \cdots, x_n].$$

Now we must check that the change of coordinates maps are smooth. If $j < i$, then

$$\phi_j \circ \phi_i^{-1}(u_1, \cdots, u_{n-1}) = \phi_j(u_1, \cdots, u_{i-1}, 1, \cdots, u_n) = \left(\frac{u_1}{u_j}, \cdots, \frac{u_{i-1}}{u_j}, \frac{1}{u_j}, \cdots, \frac{u_n}{u_j}\right),$$

which is smooth. Other computations are similar, of course. $\square$

Exercise 1.1.
Define complex projective space similarly to real projective space, and prove that it is a manifold.

Exercise 1.2.
If M and N are manifolds, then so is $M \times N$.

Exercise 1.3.
Let V be a finite-dimensional vector space over $\mathbb{R}$. Prove that V is a manifold.

1.2 Morphisms of Manifolds

We now study structure-preserving maps of manifolds.

Definition 1.10. Smooth Map
Let M and N be manifolds with atlases $\{(U_\alpha, \phi_\alpha)\}$ and $\{(V_\beta, \psi_\beta)\}$, respectively. A continuous map $f : M \to N$ is a **smooth map** (or a morphism of C^∞ manifolds) if for all α and β with

$$f^{-1}(V_\beta) \cap U_\alpha \neq \emptyset,$$

we have that the composition

$$\psi_\beta \circ f \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap f^{-1}(V_\beta)) \to \psi_\beta(V_\beta)$$

is C^∞.

Often we will write C^∞ to denote smooth. What this definition says is essentially this : Take a map, and write it out in coordinates. If it smooth in coordinates, then it is smooth. Note also that this definition will not depend on which atlas we choose. Also note a special case of this definition: $f : M \to \mathbb{R}$ is smooth if f is continuous and if for all coordinate charts $\{(U_\alpha, \phi_\alpha)\}$, $f \circ \phi_\alpha : U_\alpha \to \mathbb{R}$ is C^∞.
Given a manifold M, the collection of smooth functions $f : M \to \mathbb{R}$ will be of special importance to us, and we will denote this set as $C^\infty(M)$.

Definition 1.11. $C^\infty(M)$
$C^\infty(M)$ denotes the collection of smooth functions from M to $\mathbb{R}$.

Now, let's look at some examples of smooth maps.

Example 1.12. Take $M = \mathbb{R}^n - \{0\}$, and let $N = \mathbb{R}P^{n-1}$. Let $\pi(v) = [v]$, the quotient map that induces the topology on $\mathbb{R}P^{n-1}$. The charts on $\mathbb{R}P^{n-1}$ are the same as last time. Now, note that $\pi^{-1}(U_i) = \{v \in \mathbb{R}^n - \{0\} : v_i \neq 0\}$. To see that π is smooth, we need to check that $\phi_i \circ \pi \circ \phi^{-1} : \pi^{-1}(U_i) \to \mathbb{R}^{n-1}$ is C^∞. But note that

$$(\phi_i \circ \pi \circ \phi^{-1})(v) = \phi_i(\pi(v)) = \phi_i([v]) = \left(\frac{v_1}{v_i}, \cdots, \frac{v_n}{v_i}\right).$$

□

Example 1.13. Let $M = \mathbb{R} = N$ and $U = \mathbb{R} = V$, $\phi(x) = x$, and $\psi(x) = x^3$. Let $f : M \to N$ be the map $x \mapsto x^3$. Is this a C^∞ map?

$$(\psi \circ f \circ \phi^{-1})(x) = \psi \circ f(x^{1/3}) = \psi(x) = x,$$

which is smooth. So f is smooth.
Now let us see if the map $h(x) = x$ is smooth. The answer to this question will be no, because $\psi \circ h \circ \phi(x) = x^{1/3}$, which is not differentiable at 0. □

Example 1.14. Constant functions are smooth maps of manifolds □

The appropriate notion of "isomorphism" in differential geometry is the following one:

Definition 1.15. Diffeomorphism
A C^∞ map $f : M \to N$ is a **diffeomorphism** if f is a homeomorphism and both f and f^{-1} are C^∞ maps.

Exercise 1.4.
If M and N are manifolds, prove that $M \times N$ is diffeomorphic to $N \times M$.

Exercise 1.5.
Show that the composition of smooth maps is smooth.

Exercise 1.6.
Let $L_A : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$ be left multiplication by $A \in \mathrm{GL}(n, \mathbb{R})$. Prove that L_A is a diffeomorphism.

Exercise 1.7.
Let M be a connected manifold, and let $p, q \in M$ be any two points. Then there is a diffeomorphism of M taking p to q.
Hint: Do the case where p and q lie in the same coordinate chart first, and then utilize the fact that connected manifolds have to be path connected (why?)

At this point, we pose the following question. Given a manifold M, must there exist non-constant C^∞ functions on it? The answer is (perhaps somewhat surprisingly) yes, and the reason is the existence of so-called **bump functions**. We will investigate this question a bit and state a couple of statements without proof, mainly because the proofs of these statements are unenlightening digressions into point-set topological obscurity and should be omitted in favor of seeing some more interesting and enlightening results later on.

Definition 1.16.
If $f : X \to \mathbb{R}$ is continuous, then $\mathrm{supp}(f)$, the *support of* f, is defined to be the set $\overline{\{x \in \mathbb{R}^n : f(x) \neq 0\}}$.

Theorem 1.17.
If $K \subset \mathbb{R}^n$ is compact and $U \subset \mathbb{R}^n$ is open with $K \subset U$, then there is a smooth function on $\mathbb{R}^n$ such that:
1) $f\Big|_k \equiv 1$.
2) $0 \leq f(x) \leq 1, \forall x$.
3) supp(f)$\subset U$.

Proof. See Conlon, Theorem 2.6.1. □

1.3 Partitions of Unity

Definition 1.18. Partition of Unity
Let $\{U_\alpha\}$ be an open cover of a manifold M. A **partition of unity subordinate to** $\{U_\alpha\}$ is a collection of C^∞ functions $\{\rho_\alpha : M \to [0, 1]\}$ such that :
1) $\mathrm{supp}(\rho_\alpha) \subset U_\alpha$ for all α.
2) $\forall m \in M$, there is a neighborhood W_m of m such that $\rho_\alpha\Big|_{W_m} \neq 0$ for only finitely many α.
3) $\sum_\alpha \rho_\alpha \equiv 1$.

Partitions of unity are important for the following reason: Throughout the semester, it will be necessary to prove global existence statements about objects which we can define only in coordinate charts (i.e., locally). Loosely speaking, partitions of unity are the primary machinery for combining these individual objects defined on coordinate charts into global objects. Now,for those a bit rusty on point-set topology, "recall" that a covering of a set W is *locally finite* if for each $x \in W$, there are only finite members of the cover that intersect x. Using this definition, we can then define *paracompactness*, which is the following generalization of compactness : Let X be a topological space. A set $W \subset X$ is said to be *paracompact* provided that every open cover of W has a local finite open refinement that covers W.

The significance of *paracompactness* for us is the following:

Theorem 1.19.
Every C^∞ manifold is paracompact.

Proof. Conlon, Corollary 1.4.6. □

This proof is one stage where it is crucial that manifolds are second-countable and Hausdorff.

Theorem 1.20.
If M is a paracompact manifold then any open cover has a partition of unity subordinate to it.

Proof. Conlon, Theorem 3.5.4. □

Thus, we will always have a partition of unity when we need it, so we will assume their existence whenever necessary. Before moving on, we give a relatively simple application of these ideas. More sophisticated applications will appear later in the course.

Proposition 1.21.
Suppose that M is a manifold, and that $K \subset M$ is closed. Suppose that $U \subset M$ is open with $K \subset U$. Then there is a function $f \in C^\infty(M)$ such that :
1) $0 \leq f \leq 1$.
2) $f\Big|_K \equiv 1$.
3) supp$(f) \subset U$.

Proof. Consider the cover $\{U, M - K\}$. Let ρ_U, ρ_V be partitions of unity on U and $V = M - K$, respectively. Then $\text{supp}(\rho_U) \subset U$, and since $\rho_V\Big|_K \equiv 0$, we have that $\rho_U\Big|_K \equiv 1$. $\square$

2 Tangent Vectors and Tangent Spaces

2.1 Tangent Vectors and Tangent Spaces

An idea behind tangent vectors and tangent spaces is that we want to generalize the idea of tangent vectors to surfaces in $\mathbb{R}^n$. Since we have no notion of "tangency" for arbitrary manifolds (which may or may not sit inside Euclidean space), this may seem very difficult to do. We need a generalized notion of tangent vectors, and $\mathbb{R}^n$ does indeed provide us with a suitable, very general notion of tangency.
Namely, given a point in $\mathbb{R}^n$, there is a $1-1$ correspondence between vectors in $\mathbb{R}^n$ and directional derivatives which act on functions $f : \mathbb{R}^n \to \mathbb{R}$ that are smooth in some neighborhood of that point. To outline this correspondence, we illustrate it in $\mathbb{R}^2$ and then generalize directly from there. Accordingly, fix $a \in \mathbb{R}^2$. With a vector $v \in \mathbb{R}^n$, we associate $D_v\Big|_a$, the directional derivative of a, by

$$(D_v\Big|_a)(f) = \frac{d}{dt}\Big|_{t=0} f(a+tv),$$

or equivalently, we have the correspondence

$$v = (v_1, v_2) \mapsto \nabla \cdot v = (v_1 \frac{\partial}{\partial x}\Big|_a, v_2 \frac{\partial}{\partial y}\Big|_a).$$

Thus, if $f \in C^\infty(\mathbb{R}^2)$,

$$(\nabla \cdot v)(f) = \Big(v_1 \frac{\partial f}{\partial x}(a), v_2 \frac{\partial f}{\partial y}(a)\Big).$$

Hence, if a is a point and v is a vector, we can define a differential operator corresponding to v acting on functions $f : \mathbb{R}^2 \to \mathbb{R}$ which are smooth in a neighborhood of a. Furthermore, this space of differential operators is a vector space, and it is called the **tangent space at** a. This specific correspondence in $\mathbb{R}^n$ forms the foundation for our general notions of tangency because, as the reader will see, we can construct an analogous space of differential operators for arbitrary points on arbitrary manifolds. Indeed, as with $\mathbb{R}^2$, the tangent space for general manifolds allows us to associate a linear space of differential operators with each point on a manifold. One advantage of this idea is that we can subsequently describe manifolds and smooth maps using the more familiar language of linear algebra.
To construct the tangent space, we first need to define, in a general fashion,

what we mean by *tangent vector*. First, observe that the directional derivative in $\mathbb{R}^n$ has the following desirable properties:
(1) Let e_j be the j^{th} coordinate vector. Then $D_{e_j}\Big|_a = \frac{\partial}{\partial x_j}$.
(2)$D_v\Big|_a$ is $\mathbb{R}$-linear.
(3)$D_v\Big|_a$ obeys Leibniz's Rule : $D_v\Big|_a(fg) = f(a)D_v\Big|_a(g) + g(a)D_v\Big|_a(f)$.
Property (3) is also known as the product rule, of course. Also, any operator satisfying the above three properties is called a **derivation**.
Next, to define a tangent vector at a point a in a manifold M, we merely *define* it as any operator on $C^\infty(M)$ satisfying properties (1)-(3) above.

Definition 2.1. Tangent Vector
Let M be a manifold with $a \in M$. A **tangent vector** to M at a is a map $v : C^\infty(M) \to \mathbb{R}$ such that :
(1) v is $\mathbb{R}$-linear, and
(2) $v(fg) = f(a)v(g) + g(a)v(f)$.

We would like to show the following things :
(1) Tangent vectors always exist;
(2) Tangent vectors form an $\mathbb{R}$-vector space (denoted by T_aM) whose dimension is the same as the dimension of M, and
(3) The coordinate charts define a basis for T_aM.
We will do this in several steps. The first two propositions are relatively simple, so we'll do them first. The third property will require some more machinery, however.

Proposition 2.2.
Tangent vectors always exist.

Proof. Let M be a manifold and $a \in M$. Suppose that (U, ϕ) is a coordinate chart with $a \in U$. Define

$$\frac{\partial}{\partial x_i}(f)\Big|_a = \frac{\partial}{\partial r_i}\Big|_{\phi(a)}(f \circ \phi^{-1}),$$

where $f \in C^\infty(M)$. Then $\frac{\partial}{\partial x_i}\Big|_a$ is a tangent vector at a. □

The only thing we did in the previous proposition is define the notion of a partial derivative on a manifold in terms of the derivative of a function from $\mathbb{R}^n \to \mathbb{R}$, and because we know how such derivatives act, it is clear from our definition of tangent vectors that we actually have defined a tangent vector.

Proposition 2.3.
Let M be a manifold, with $a \in M$. Then T_aM is an $\mathbb{R}$-vector space.

Proof. If v and w are two tangent vectors and λ, μ are two real numbers, define $(\lambda v + \mu w)(f) = \lambda v(f) + \mu w(f)$. By definition, this makes T_aM into a vector space. $\square$

Definition 2.4. Tangent Space
Let M be a manifold, and let $a \in M$. Then T_aM is called the **tangent space to** M **at** a.

Note that the above definition makes sense since we know that our ordinary differential operators on $\mathbb{R}$ are $\mathbb{R}$-linear. Now we state the main theorem of the section, which we will prove via a couple of lemmas.

Theorem 2.5.
Suppose that $(U, \phi = (x_1, \cdots, x_n))$ is a coordinate chart on M. Then $\{\frac{\partial}{\partial x_j}\}$ forms a basis for T_aM, and for any $v \in T_aM$,

$$v = \sum_i v(x_i) \frac{\partial}{\partial x_i}.$$

That is, the coordinate charts help us define a basis for T_aM, and a tangent vector is completely determined by its action on the coordinate functions. In order for this to make sense, first we need show that our definition of tangent vectors is local, in the sense that we can locally identify T_aU with T_aM if (U, ϕ) is a coordinate chart on M. This is of particular importance since global charts may not always be defined, while local charts are always defined.

Lemma 2.6.
Suppose that $v \in T_aM$. Then
(1) If $f, g \in C^\infty(M)$ are such that $f = g$ in a neighborhood U of a, then $v(f) = v(g)$.
(2) If h is constant on a neighborhood U of a, then $v(h) = 0$.

Proof. (1) Suppose that $f\Big|_U = g\Big|_U$. As v is $\mathbb{R}$-linear, it is enough to show that $v(f - g) = 0$. Pick $\rho \in C^\infty(M)$ with $0 \leq \rho \leq 1$, $\rho \equiv 1$ "near" a, and $\text{supp}(\rho) \subset U$. Now, we have that $\rho(f - g) = 0$ on all of M by construction. Furthermore, because v is linear, $v(0) = 0$, so that

$$0 = v(\rho(f - g)) = v(\rho)(f - g)(a) + \rho(a)v(f - g) = v(f - g).$$

(2) Let 1 denote the constant function 1. By linearity, we have $v(1 \cdot 1) = 2v(1)$, so that $v(1) = 0$. Now, for any constant function c, $v(c) = c \cdot v(1) = 0$. By part (1), we are done. $\square$

Lemma 2.7.
Let M be a manifold, $U \subset M$ be an open set, and $a \in U$. Then the map $T_aU \to T_aM$ given by $v \mapsto (f \mapsto v(f\big|U))$ is an isomorphism.

Proof. By the previous lemma, the map is well-defined, and it is linear by definition of tangent vectors. To see that it is an isomorphism, pick a bump function $\rho \in C^\infty(U)$ whose support is in U and which is identically 1 "near" a. Then the map from $T_aM \to T_aU$ given by $v \mapsto (f \mapsto v(\rho f))$ is an inverse. $\square$

Lemma 2.8.
Suppose that $h \in C^\infty(B_R(0))$. Then $h(r) = h(0) + \sum r_i h_i(r)$, where $h_i(0) = \frac{\partial h}{\partial r_i}(0)$.

Proof.

$$h(r) - h(0) = \int_0^1 (\frac{d}{dt}h(tr))dt = \int_0^1 (\sum r_i \frac{\partial h}{\partial r_i}(tr))dt = \sum r_i \int_0^1 \frac{\partial h}{\partial r_i}(tr)dt.$$

$\square$

Now we are in a position to prove that the coordinate differentials form a basis for T_aM.

Proof. We want to show that if $(U, \phi = (x_1, \cdots, x_n))$ is a coordinate chart on M, then $\frac{\partial}{\partial x_i}\Big|_a = \frac{\partial}{\partial r_i}\Big|_{\phi(a)}(f \circ \phi^{-1})$ form a basis of T_aM. Note that $\frac{\partial}{\partial x_i}\Big|_a(x_j) = \frac{\partial}{\partial r_i}\Big|_{\phi(a)}(r_j \circ \phi^{-1})$. But $x_j = r_j \circ \phi$, so this expression reduces to $\frac{\partial}{\partial r_i}(x_j)\Big|_a = \delta_{ij}$. This then implies that the $\frac{\partial}{\partial x_j}$'s are linearly independent.
Now, we argue that for all $v \in T_aU$ and all $f \in C^\infty(U)$,

$$v(f) = \sum v(x_i) \frac{\partial}{\partial x_i}\Big|_a (f).$$

WLOG, we may assume that $\phi(a) = 0$ in $\mathbb{R}^n$ (we always may translate our charts back to the origin). Furthermore, assume WLOG that $\phi[U] = B_R(0)$. Then we have

$$(f \circ \phi^{-1})(r) = (f \circ \phi^{-1})(0) + \sum r_i h_i(r),$$

by the previous lemma, where $h_i(0) = \frac{\partial}{\partial r_i}(f \circ \phi^{-1})\Big|_0$. Thus,

$$f(x) = f(a) + \sum x_i \cdot f_i(x),$$

where

$$f_i(a) = \frac{\partial}{\partial r_i}(f \circ \phi^{-1})(0) = \frac{\partial}{\partial x_i}\Big|_a (f),$$

for all $x \in U$. Hence, for any $v \in T_aM$, we have

$$\begin{aligned} v(f) &= v(f(a) + \sum x_i f_i) \\ &= \sum x_i(a) v(f_i) \sum v(x_i) f_i(a) \\ &= \sum v(x_i) f_i(a) \\ &= \sum v(x_i) \frac{\partial}{\partial x_i}\Big|_a (f). \end{aligned}$$

□

Note the following special case of this theorem: Namely, if $(U, \psi = (y_1, \cdots, y_n))$ is another coordinate chart, then

$$\frac{\partial}{\partial y_j}\Big|_a = \sum \frac{\partial}{\partial y_j}\Big|_a (x_i) \cdot \frac{\partial}{\partial x_i}\Big|_a.$$

This is the change of coordinates formula, where

$$\frac{\partial}{\partial y_j}\Big|_a (x_i) = \frac{\partial}{\partial r_i}(x_i \circ \psi^{-1}) = \frac{\partial}{\partial r_i}(r_j \circ (\phi \circ \psi^{-1})),$$

the Jacobian of $\phi \circ \psi^{-1}$. So if we have two coordinate charts on U, then the change of coordinates formula is given by the Jacobian of the composition. As a notational convention, $\{\frac{\partial}{\partial r_i}\}_i$ is a basis of $T_a\mathbb{R}^n$ for all a. We identify $\mathbb{R}^n$ with $T_a\mathbb{R}^n$, since there is a natural isomorphism between the two spaces. Also, from now on, we will use $\{r_i\}$ to denote the standard (global) coordinates on $\mathbb{R}^n$.

2.2 Differentials

With each point on a manifold, we have associated a linear space whose dimension as a vector space over $\mathbb{R}$ is equal to the dimension of the manifold. But what can we say about mappings between these linear spaces? As it turns out, every smooth map of manifolds will induce linear maps between tangent spaces. Furthermore, in the familiar case of mappings from $\mathbb{R}^n$ to $\mathbb{R}^m$, this induced linear map will turn out to be the derivative; i.e., the $m \times n$ matrix of first partial derivatives evaluated at the appropriate point. We now give the necessary definitions.

Definition 2.9. Differential
Suppose that $F : M \to N$ is C^∞. Then for all $a \in M$ and all $f \in C^\infty(M)$, we define the **differential of F**

$$dF : T_aM \to T_{F(a)}N$$

by

$$(dF_a(v))(f) = v(f \circ F).$$

Note that since tangent vectors are $\mathbb{R}$-linear, the differential map is also $\mathbb{R}$-linear. Thus, the differential map is a homomorphism of vector spaces, or a linear transformation. Furthermore, since a tangent vector is completely determined by its action on coordinate functions (which we proved in the last section), we can compute differentials of mappings, which we will do in a couple of examples below. First, however, we will consider two special cases of the differential map; namely, when $M = \mathbb{R}$ and when $N = \mathbb{R}$.
Case (1): $N = \mathbb{R}$.
In this case, $dF_a : T_aM \to T_{F(a)}\mathbb{R}$, which we identify with R; that is, $dF_a \in T_aM^*$, the dual space of T_aM. More precisely,

$$(dF_a(v))(r) = v(r \circ F) = v(F);$$

that is, $dF_a(v) = v(F)$. Note that $v = \sum v(x_i)\frac{\partial}{\partial x_i} = \sum dx_i(v)\frac{\partial}{\partial x_i}$, which says that $\{dx_i\}$ forms a basis for the dual space T_aM^*.
Case (2): $M = \mathbb{R}$.
In this case, $F : \mathbb{R} \to N$ is a curve. Then $(dF)_t(\frac{d}{dr}) = F'(t)$, the tangent vector to a curve at $F(t)$.
Fortunately, differentials obey a nice composition rule: the familiar **chain rule** from calculus.

Theorem 2.10. Chain Rule
If $F : X \to Y$ and $H : Y \to Z$ are smooth maps of manifolds, then

$$d(H \circ F)_a = dH_{F(a)} \circ dF_a.$$

Proof. Fix $a \in X$, $v \in T_aX$, and $f \in C^\infty(Z)$. Then

$$(d(H \circ F)_a(v))(f) = v(f \circ (H \circ F))$$

$$\begin{aligned} &= v((f \circ H) \circ F) \\ &= (dF_a(v))(f \circ H) \\ &= (dH_{F(a)}(dF_a(v)))(f). \end{aligned}$$

□

Example 2.11.
Suppose that $f : \mathbb{R}^m \to \mathbb{R}^n$ is smooth. We have defined $df_x : T_xM \to T_{f(x)}N$; it is natural to ask how this relates to the derivative map $Df_x : \mathbb{R}^m \to \mathbb{R}^n$, which is the matrix of partials evaluated at the point $x \in \mathbb{R}^m$. We guess that these maps are the same (after we identify $T_x\mathbb{R}^m$ with $\mathbb{R}^m$ and $T_{f(x)}\mathbb{R}^n$ with $\mathbb{R}^n$). To see this, we'll simply compute

$$df_x(\frac{\partial}{\partial r_j})\Big|_x.$$

Writing a vector $w \in T_x\mathbb{R}^m$ out in coordinates, we have

$$w = \sum w(s_i)\frac{\partial}{\partial s_i}\Big|_x.$$

Then

$$\begin{aligned} (df_x(\frac{\partial}{\partial r_j}\Big|_x))(s_i) &= \frac{\partial}{\partial r_j}\Big|_x(s_i \circ f) \\ &= \frac{\partial}{\partial r_j}\Big|_x(f_i) \\ &= \frac{\partial f_i}{\partial r_j}(x) \\ &= Df_x(e_j). \end{aligned}$$

In other words, the differential at a point *is* the Jacobian matrix of first partial derivatives evaluated at that point. □

Exercise 2.1.
If M and N are manifolds, then $T_{m,n}(M \times N) = T_mM \times T_nN$ for all $m \in M$ and $n \in N$.

Exercise 2.2.
If $\phi : M \to N$ is a diffeomorphism, then

$$d\phi_a : T_aM \to T_{\phi(a)}N$$

is an isomorphism.

Exercise 2.3.
Suppose that $\gamma : \mathbb{R} \to \mathbb{R}^n$. Show that

$$d\gamma(\frac{d}{dt}) = \sum_i \gamma^{i\prime}(t) \frac{\partial}{\partial r_i}.$$

Exercise 2.4.
For any tangent vector $v \in T_m M$, there is a curve $\gamma : (a, b) \to M$ with $\gamma(0) = m$ and $d\gamma_0(\frac{d}{dt}) = v$.

2.3 The Tangent Bundle

Definition 2.12. Tangent Bundle
The **tangent bundle** TM to a manifold M is

$$TM = \bigcup_{a \in M} T_a M.$$

We want to show that TM itself is a manifold. Strictly speaking, we first should specify a topology on TM, but let us find candidates for charts on TM instead. To do this, begin with a coordinate chart $(U, \phi = (x_1, \cdots, x_n))$ on M. Let $U^* = TU$. Note that there is a map $\pi : TM \to M$ that sends $v \in T_a M$ to a, which is called the **bundle projection**. Define

$$\phi^* = (x_1 \circ \pi, \cdots, x_n \circ \pi, dx_1, \cdots, dx_n).$$

Now, if $\{(U_\alpha, \phi_\alpha)\}$ is an atlas on M, we claim that $\{(U^*_\alpha, \phi^*_\alpha)\}$ is an atlas on TM.
To see why this claim is true, let $(U, \phi = (x_1, \cdots, x_n))$ and $(V, \psi = (y_1, \cdots, y_n))$ be two coordinate charts on M with $U \cap V \neq \emptyset$. Then $T(U \cap V) = TU \cap TV$, and $\phi^* = (x_1, \cdots, x_n, v_1, \cdots, v_n)$. Also, $\psi^* = (y_1, \cdots, y_n, w_1, \cdots, w_n)$ are candidates for charts on U^* and V^*. Now let us compute
$(\psi^* \circ \phi^{*-1})(r_1, \cdots, r_n, u_1, \cdots, u_n)$, to see whether or not change of coordinates is smooth. First, note that

$$\phi^{*-1}(r_1, \cdots, r_n, u_1, \cdots, u_n) = \sum u_i \frac{\partial}{\partial x_i}\Big|_{\phi^{-1}(r_1, \cdots, r_n)} \in T_{\phi^{-1}(r_1, \cdots, r_n)} M.$$

So

$$\psi^*(\sum u_i \frac{\partial}{\partial x_i}\Big|_{\phi^{-1}(r_1, \cdots, r_n)}) = \psi(\phi^{-1}(r_1, \cdots, r_n), dy_1(\sum u_i \frac{\partial}{\partial x_i}), \cdots, dy_n(\sum u_i \frac{\partial}{\partial x_i})).$$

But
$$dy_j(\sum u_i \frac{\partial}{\partial x_i}) = \sum u_i(\frac{\partial}{\partial x_i}(y_j)) = \frac{\partial}{\partial r_j}(r_j(\psi \circ \phi^{-1})),$$
a piece of the Jacobian. Thus, under this composition map, we have that $(r_1, \cdots, r_n, u_1, \cdots, u_n) \mapsto ((\psi \circ \phi^{-1}(r_1, \cdots, r_n)), (\sum_i A_{ij} u_i)_j)$, where $A_{ij} = \frac{\partial}{\partial r_j}(r_j \circ (\psi \circ \phi^{-1}))$. If we induce the topology of $\phi^* : TU \to \phi(U) \times \mathbb{R}^n$ so that ϕ is a homeomorphism, we then have that our functions ϕ^* are homeomorphisms, and we have just shown that change of coordinates is smooth. Thus, to finish the proof, we need the following result, which we should prove, but won't:

Proposition 2.13. *If M is Hausdorff and second countable, so is TM.*

Thus, it follows that TM is a manifold, and in fact, we have shown that if M is n-dimensional, TM is a $2n$-dimensional manifold. As a matter of notation, we write $(m, v) \in TM$ for $v \in T_m(M)$.

Exercise 2.5.
Prove that the map $\pi : TM \to M$ is smooth and that the differential $d\pi_v : T_v(TM) \to T_{\pi(v)}M$ is surjective.

2.4 The Cotangent Bundle

By analogy with the tangent bundle, we can define the **cotangent bundle** as
$$T^*M = \coprod_{a \in M} T_a^*M.$$
Recall from earlier that if $(U, \phi = (x_1, \ldots, x_n))$ is a coordinate chart on M, then for $a \in U$, $\{dx_1, \ldots, dx_n\}$ forms a basis for T_a^*M, the **cotangent space of** M **at** a. One can think of T_a^*M as the dual space of the vector space T_aM. Once again, we write (a, v) for the element $v \in T_a^*M$. We now have the following theorem:

Theorem 2.14.
If M is an n-dimensional manifold, then T^*M is a $2n$-dimensional manifold. In fact, if $(U, \phi = (x_1, \ldots, x_n))$ is a coordinate chart on M, then
$$(U \times \mathbb{R}^n, \phi^* = (x_1, \ldots, x_n, \frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}))$$
defines a coordinate chart on T^*M. Moreover, the projection map $\pi : T^*M \to M$ given by $\pi(a, v) = a$ is smooth.

Proof. Since we have defined charts, we need only check that change of coordinates is smooth, and this computation is quite similar to the change of coordinates computation we did on the tangent bundle. Namely, if $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ are two local coordinate systems on M, then the change of coordinates map $\psi : U \times \mathbb{R}^n \to U \times \mathbb{R}^n$ is given by

$$(q, (\eta_1, \ldots, \eta_n)) \mapsto (q, \ldots, \sum_j \eta_j dx_j(\frac{\partial}{\partial y_i}), \ldots).$$

□

Remark 2.15.
Suppose $f : M \to \mathbb{R}$ is smooth. Then at each point $x \in M$, df_x is an element of the dual of T_xM; that is, we may view df as a map from $M \to T^*M$. This point of view will be important when we study differential forms later on in the semester.

Exercise 2.6.
Suppose that $f : M \to N$ is a diffeomorphism. Show that f lifts to a diffeomorphism $\tilde{f} : T^*M \to T^*N$.
Hint: For every point $x \in M$, we have $df_x : T_xM \to T_{f(x)}N$. Since f is a diffeomorphism, we get $(df_x)^{-1} : T_{f(x)}N \to T_xM$, and by taking the transpose, we get $(df_x^{-1})^T : T_x^*M \to T_{f(x)}^*N$. Define

$$\tilde{f}(x, \eta) = (f(x), (df_x^{-1})^T \eta),$$

where $x = \pi(\eta)$.

2.5 Vector Fields

Now we define the notion of a vector field on a manifold.

Definition 2.16. Vector Field
A **vector field**X in a manifold M is a C^∞ map

$$X : M \to TM$$

such that

$$(\pi \circ X)(a) = a \; \forall a \in M.$$

Equivalently, we have that $X(a) \in T_aM \; \forall a \in M$.

In other words, we define a vector field to be a section of the tangent bundle. The following proposition gives us an alternate characterization of vector fields:

Proposition 2.17.
The following notion of a vector field is equivalent to the one above: A vector field X on a manifold M is a map $X : C^\infty(M) \to C^\infty(M)$ such that

$$X(fg) = fX(g) + gX(f) \;\forall f, g \in C^\infty(M).$$

Exercise 2.7.
Prove the previous proposition.

As a matter of notation, vector fields on a manifold M are usually represented by $\chi(M)$ or $\Gamma(TM)$. Note that $\Gamma(TM)$ is both an $\mathbb{R}$-vector space and a $C^\infty(M)$-module.
Now, it is not the case that $X(Y)$ defines a vector field, but it may be a surprising fact that the "multiplication" defined by $[X, Y] = X(Y) - Y(X)$ does define a vector field. This multiplication of vector fields is called the **Lie bracket of X and Y**. Its properties are summed up in the following proposition.

Definition 2.18.
A Lie algebra is a real vector space V together with a multiplication $[\cdot, \cdot] : V \times V \to V$ called the **bracket** which satisfies the following properties for all $X, Y, Z \in V$:
(1) The bracket is bilinear over $\mathbb{R}$;
(2) $[X, Y] = -[Y, X]$;
(3) Jacobi Identity: $[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]]$.

Proposition 2.19.
$\Gamma(TM)$ equipped with $[\cdot, \cdot]$ is a Lie algebra.

Proof. Let $X, Y, Z \in \Gamma(TM)$. Using the previous proposition (i.e., the second characterization of vector fields) it is not hard to check that $[X, Y]$ is a vector field. In addition, bilinearity and skew-commutativity are immediate properties, so we'll verify the Jacobi identity. Let $f \in C^\infty(M)$; then

$$\begin{aligned} [X, [Y, Z]](f) &= X([Y, Z](f)) - [Y, Z](X(f)) \\ &= X(Y(Z(f))) - X(Z(Y(f))) - Y(Z(X(f))) + Z(Y(X(f))). \end{aligned}$$

Permuting cyclically and adding establishes the identity. □

Exercise 2.8.
Let X and Y be vector fields on M. Then $[X, Y]$ is also a vector field on M.

Exercise 2.9.
Compute
$$[x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}, x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}],$$
the Lie bracket of two vector fields on $\mathbb{R}^2$.

3 Submanifolds and Implicit Function Theorem

3.1 The Inverse and Implicit Function Theorems

Before defining the notion of a submanifold, we take time to state two theorems which are of fundamental importance in differential geometry. The first of these is the inverse function theorem, which provides a "local converse" to the earlier exercise that the differential of a diffeomorphism is an isomorphism.

Theorem 3.1. Inverse Function Theorem
Suppose that $F : V \to W$ is smooth, where V and W are Banach spaces, and suppose that dF_{x_0} is an isomorphism. Then there is a neighborhood U of x_0 such that $F\Big|_U : U \to F(U)$ is a diffeomorphism.

Immediately, we then have the following corollary:

Corollary 3.1.1.
Suppose that M and N are manifolds and that $f : M \to N$ is smooth. If $df_x : T_xM \to T_{f(x)}N$ is an isomorphism, then there is some neighborhood U of x such that $f : U \to f[U]$ is a diffeomorphism.

Proof. Let (U, ϕ) be a coordinate chart on M with $x \in U$ and (ψ, V) be a coordinate chart on N such that $f(x) \in V$. WLOG, we may assume that $\phi[U] = \mathbb{R}^m$, $\psi[V] = \mathbb{R}^n$, $\phi(x) = 0$, $\psi(f(x)) = 0$. Then the following diagram commutes:

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[d, "\phi"'] & V \arrow[d, "\psi"] \\ \mathbb{R}^m \arrow[r, "\psi\circ f\circ\phi^{-1}"'] & \mathbb{R}^n \end{tikzcd}$$

Since df_x is an isomorphism, so is $d(\psi \circ f \circ \phi^{-1})_{\phi(x)}$; apply the previous theorem to invert this map in some small neighborhood of $0 \in \mathbb{R}^n$. Defining f^{-1} on some neighborhood of $f(x)$ is now easy to do. $\square$

Usually, the inverse function theorem is proved using contraction mappings on metric spaces. If you're curious, there's a proof in Rudin's undergraduate analysis book.
We now state a version of the implicit function theorem, which can be proved as a consequence of the inverse function theorem.

Theorem 3.2. Implicit Function Theorem
Let $p = (p_1, \dots, p_n) \in \mathbb{R}^n$. Given $f_1, \dots, f_k \in C^\infty(p)$ with

$$\det(\frac{\partial f_i}{\partial x_j}(p))_{1 \le i,j \le k} \neq 0,$$

there exist functions $y_1, \dots, y_k \in C^\infty((p_{k+1}, \dots, p_n))$ such that in a neighborhood of p in $\mathbb{R}^n$,

$$f_j(x_1, \dots, x_n) = 0, \ \forall j$$

if and only if

$$x_j = y_j(x_{k+1}, \dots, x_n), \ 1 \le j \le k.$$

Now for the definition of a submanifold.

Definition 3.3. Submanifold
Let M be an m-dimensional manifold. A subset $N \subset M$ is an n-dimensional **submanifold** if for every point $x \in N$, there is a coordinate chart $(U, \phi = (x_1, \cdots, x_m))$ with $x \in U$ such that

$$\phi(U \cap N) = (\phi(U)) \cap (\mathbb{R}^n \times 0).$$

That is, for all $a \in N$, $\phi(a) = (x_1(a), \cdots, x_n(a), 0, \cdots, 0)$. Such charts are said to be **adapted to N**.

One of the main features of this definition is that any submanifold is itself a manifold *in the subspace topology.* That is, if N is a submanifold of M, then the topologies of N and M are compatible in the sense that N inherits its topology as a subspace of M. Recall that the subspace topology of M in N is defined as follows: a subset $U \subset N$ is open in the subspace topology if and only if $U = O \cap N$ for some open subset of M.
There are more general definitions of submanifolds, and for this reason, our notion of a submanifold is sometimes called a *regular submanifold.* Since a submanifold behaves somewhat like a projection of a larger space onto a smaller space locally, the following proposition should not be surprising.

Proposition 3.4.
Let N be a submanifold of a manifold M, and let $\iota : N \to M$ be the inclusion. Then ι is smooth, and its differential is injective.

Proof. Let ϕ denote an adapted coordinate chart on N and let ψ denote a local coordinate chart on M. Then $\psi \circ i \circ \phi^{-1} : \mathbb{R}^m \to \mathbb{R}^n$ is given by

$$(x_1, x_2, \dots, x_m) = (x_1, \dots, x_m, 0, \dots, 0),$$

certainly a smooth map.
To see injectivity, note that if $\{x_1, \dots, x_n\}$ are local coordinates on M and $\{x_1, \dots, x_m\}$ are corresponding adapted coordinates in some neighborhood of $x \in N$, then

$$di_x(\frac{\partial}{\partial x_i}) = \frac{\partial}{\partial x_i}.$$

□

Example 3.5.
Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth map. Then the graph of f, Γ is a submanifold of $\mathbb{R}^{n+1}$. To see this, we need to produce a map ϕ from some open set U in $\mathbb{R}^{n+1}$ to $\mathbb{R}^{n+1}$ which sends $U \cap \Gamma$ to $\mathbb{R}^n \times 0$. Take $\phi(x, y) = (x, y - f(x))$; it works!

3.2 Regular Values

One way by which we can prove that certain subsets of manifolds are themselves submanifolds is through the notion of a **regular value**. We now give the definition and prove the main theorem involving regular values.

Definition 3.6. Regular Value
Suppose $f : M \to N$ is smooth. Then $y_0 \in N$ is a **regular value** of f if for all $x \in f^{-1}(y_0)$, $df_x : T_xM \to T_{y_0}N$ is surjective.

Remark 3.7. Note that if $f^{-1}(y_0) = \emptyset$, it is a regular value of f.

The above remark may seem silly, but it is important for the following reason:

Theorem 3.8. Sard's Theorem
Let $f : M \to N$ be a smooth map. Then the set of regular values of f is dense in M (and in fact has measure 0).

We will not prove the above theorem but instead focus on the geometric properties of regular values. Here is the most important of these for our purposes.

Theorem 3.9.
If y_0 is a regular value of $f : M \to N$, then $f^{-1}(y_0)$ is a submanifold of M of dimension $\dim(M) - \dim(N)$.

Proof. Let $Z = f^{-1}(y_0)$. Pick $m \in Z$, and let $(U, \phi = (x_1, \dots, x_m))$ be a coordinate chart on M with $m \in U$. Without loss of generality, we may assume that $\phi(m) = 0$. Furthermore, let $(V, \psi = (y_1, \dots, y_n))$ be a coordinate chart near y_0. Define $h = \psi \circ f \circ \phi^{-1}$; it is enough to show that $h^{-1}(0) \cap \phi(U)$ is a submanifold of $\phi(U)$ near 0.
Now, by the chain rule, $dh = d\psi \circ df \circ d\phi^{-1}$, and it's a surjective map because $d\psi$ and $d\phi^{-1}$ are invertible linear maps and since df is a surjection. Therefore, we may assume that $M = R^m$, $N = R^n$, $f : M \to N$ with $f(0) = 0$, and $df_x : \mathbb{R}^m \to \mathbb{R}^n$ is surjective for all x.
Let $V = \ker df_0$. Then $\mathbb{R}^m = V \oplus W$, and $df_0 : W \to \mathbb{R}^n$ is an isomorphism. Next, define $H : V \times W \to V \times \mathbb{R}^n$ by $H(v, w) = (v, f(v, w))$. Note that the differential of H at the origin has the following form:

$$\begin{pmatrix} I & 0 \\ df_0\big|_V & df_0\big|_W \end{pmatrix}$$

In particular, this says that $dH_{(0,0)}$ is an isomorphism, so that by the inverse function theorem, there is a neighborhood U_0 of $(0, 0)$ in $V \times W$ such that $H(U_0) \subset V \times \mathbb{R}^n$ is open and $H : U_0 \to H(U_0)$ is an isomorphism. Furthermore, note that $f^{-1}(0) \cap U_0 = \{(v, w) \in U_0 : H(v, w) = (v, 0)\}$. Accordingly, take (U_0, H) as a coordinate chart, for then $H(f^{-1}(0) \cap U_0) = U_0 \cap (V \times 0)$. □

Corollary 3.9.1.
If a is a regular value of $f : M \to N$, then $\dim f^{-1}(a) = \dim(M) - \dim(N)$.

Corollary 3.9.2.
Suppose that y_0 is a regular value of $f : M \to N$ and $f^{-1}(y_0) \neq \emptyset$. Then for all $m \in f^{-1}(y_0)$, $T_m f^{-1}(y_0) = \ker(df_m)$.

Proof. If A is a manifold and $(q, v) \in TQ$, then there is a curve $\gamma : (a, b) \to Q$ such that $\gamma(0) = q$ and $d\gamma(\frac{d}{dt}) = v$, by a previous exercise. Now, observe that $\dim T_m f^{-1}(y_0) = \dim \ker df_m$, so it is enough to prove that $T_m f^{-1}(y_0) \subset \ker df_m$. Let $v \in T_m f^{-1}(y_0)$. Then $d(f \circ \gamma)_0(\frac{d}{dt}) = 0$, since $f \circ \gamma$ is a constant map. Then by the chain rule, $d(f \circ \gamma)_0(\frac{d}{dt}) = df_{\gamma(0)}(d\gamma_0(\frac{d}{dt})) = df_m(v)$. □

Now let's look at some applications of this.

Example 3.10.
Let $f : \mathbb{R}^n \to \mathbb{R}$ be given by $x \mapsto ||x||$. Then $df_x = (2x_i)_i$, which means that df_x is onto for all x except the origin. In particular, this means that 1 is a regular value of f, which shows, via the previous theorem, that S^{n-1} is a submanifold of $\mathbb{R}^n$. □

Example 3.11.
Recall that $G = \mathrm{GL}(n, \mathbb{R})$ is a manifold of dimension n^2. Consider the determinant map from G to $\mathbb{R}$; it is smooth since it is a polynomial mapping. Furthermore, 1 is a regular value of det. To see this, let A have determinant 1, and note that the differential $\det_A$ is a $1 \times n$ of partial derivatives, not all of which can be 0. Thus, $\mathrm{SL}(n, \mathbb{R})$ is a submanifold of G, and it has dimension $n^2 - 1$. $\square$

Exercise 3.1.
Show that $\mathrm{O}(n)$, the set of all $n \times n$ orthogonal matrices, is a submanifold of $\mathrm{GL}(n, \mathbb{R})$.
Hint: Consider the map $f : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{Sym}(n, \mathbb{R})$ given by $A \mapsto AA^T$. Show that I is a regular value.

3.3 Transversality

We now explore another property related to submanifolds, which is actually a generalization of regular values.

Definition 3.12. Transversality
Let $F : M \to N$ be a C^∞ map and $Z \subset N$ a submanifold. F is **transverse to** Z if for every $z \in Z$ and any $m \in F^{-1}(z)$, we have

$$T_z Z + dF_m(T_m M) = T_z N.$$

Thinking in terms of transverse things, the above definition says that $T_z Z$ contains the orthogonal complement of the differential image (and vice versa). That is, "transverse" really does make sense here, and we can think of it as the opposite of tangency.

Example 3.13.
Suppose that $Z = \{y_0\}$. Then Z is transverse to F if and only if y_0 is a regular value of F. $\square$

Example 3.14.
Take $M = N = \mathbb{R}^2$. Consider $F : M \to N$ given by $F(x, y) = (x, x^2)$. Then F is transverse to $0 \times \mathbb{R}$, but it is not transverse to $\mathbb{R} \times 0$. $\square$

The big theorem regarding transversality is the following one:

Theorem 3.15.
If $F : M \to N$ is transverse to Z, then $F^{-1}(Z)$ is a submanifold of M. Moreover, for all $m \in F^{-1}(Z)$,

$$T_m(F^{-1}(Z)) = (dF_m)^{-1}(T_{F(m)} Z).$$

Thus, $\dim(M) - \dim(F^{-1}(Z)) = \dim(N) - \dim(Z)$.

Proof. First, let us consider the special case that $N = \mathbb{R}^n$, $Z = \mathbb{R}^k \times 0 \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$. Let $\pi : \mathbb{R}^k \times \mathbb{R}^{n-k} \to \mathbb{R}^{n-k}$ denote the canonical projection map. Now, note the following:
(1) $(\pi \circ F)^{-1}(0) = F^{-1}(Z)$.
(2) $d(\pi \circ F)_m(T_m M) = d\pi_{F(m)}(dF(T_m M)) = d\pi_{F(m)}(dF_m(T_M) + T_{F(m)}Z) = d\pi_{F(m)}(\mathbb{R}^n) = \mathbb{R}^{n-k}$.
Thus, $d(\pi \circ F)_m(T_m M)$ is surjective; that is, 0 is a regular value of $(\pi \circ F)$, and $(\pi \circ F)^{-1}(0) = F^{-1}(Z)$ is a submanifold of M. Moreover, $T_M F^{-1}(Z) = \ker d(\pi \circ F)_m = \ker d\pi_{F(m)} \circ dF_m = (dF_m)^{-1}(\ker d\pi) = (dF_m)^{-1}(T_{F(m)}Z)$.
Finally, $\dim F^{-1}(Z) = \dim \ker d\pi \circ F_m = \dim M - \dim \mathbb{R}^{n-k}$, since $(d\pi \circ F)_m$ is surjective. Then we have $\dim M - \dim R^{n-k} = \dim M - \dim N - \dim Z$.
Now, to see why the special case implies the general case, simply note that for all $z \in Z$, there is a coordinate chart $(V, \psi = (x_1, \ldots, x_n))$ such that $\psi(Z) = \psi(V) \cap (\mathbb{R}^k \times 0)$. □

Here are a couple of interesting examples of this theorem in action.

Example 3.16.
Consider two surfaces S_1 and S_2 in $\mathbb{R}^3$ such that $T_x S_1 \neq T_x S_2$ for every $x \in S_1 \cap S_2$. Let $F = \iota_1 : S_1 \to \mathbb{R}^3$, the inclusion map. Note that $dF_x(T_x S_1) = T_x S_1$, and as the tangent spaces of these surfaces are not equal, we must have $T_x S_1 + T_x S_2 = \mathbb{R}^3$, for all $x \in S_1 \cap S_2$. Thus, F is transverse to S_2, which says that $F^{-1}(S_2) = S_1 \cap S_2$ is a submanifold of S_1 of dimension 1, by the previous theorem. □

Example 3.17.
Let $Z \subset N$ be a submanifold. Consider $\pi : TN \to N$, which we know to be surjective at every point in N by a previous exercise. In other words, we have that π is transverse to Z, which says that $\pi^{-1}(Z) \subset TN$ is a submanifold of TN. That is, $TN\Big|_Z$ is a submanifold of TN. □

3.4 Embeddings, Immersions, and Rank

Definition 3.18. Immersion
A smooth map $f : Z \to M$ is an **immersion** if its differential is always $1 - 1$.

$f[Z]$ is often called an *immersed submanifold* of M. Note that every submanifold is an immersed submanifold, where one takes the immersion to be the inclusion map.

Here's a concept with which we will not do that much, but we'll go ahead and define it nevertheless.

Definition 3.19. Submersion
A map $f : M \to N$ is called a **submersion** if its differential at every point is surjective.

Note that if f is a submersion, then every point of N is a regular value of f; thus, $f^{-1}(n)$ is a subamnifold of M for every $n \in N$.

Definition 3.20. Embedding
A smooth map $f : Z \to M$ is an **embedding** if $Z \subset N$ is a submanifold and $f : Z \to f[Z]$ is a diffeomorphism.

The image $f[Z]$ is said to be an *embedded submanifold*. An equivalent way to define an embedding is to say that it is a 1-1 immersion that is a homeomorphism into M. The point here is that an embedded submanifold inherits its topology from the larger space. An embedded submanifold is a submanifold when the embedding is given by the inclusion map.
Thus, examples of embeddings are furnished by submanifolds and their inclusion maps, so let's look at some examples of immersions that are not embeddings to really illustrate the difference between the two concepts.

Example 3.21.
Let $f : S^1 \to \mathbb{C}$ be given by $f(z) = z^2$. Now, $df_z = 2z$, and so it is easy to see that f is an immersion. Furthermore, f is locally 1-1 , but it is not globally injective (since, it is, in fact, a 2-1 mapping). So f is not an embedding. $\square$

Example 3.22.
Let A represent a (non-closed) figure 8 in the plane, which we can consider to be the image of an interval under some smooth map from $\mathbb{R}$ into the plane. Let us call this map g. Then is an example of a 1-1 immersion which is not an embedding; in other words, A is an embedded submanifold which is not an immersed submanifold. $\square$

Example 3.23.
Consider the following map of $\mathbb{R}$ into the torus (which we identify with $S^1 \times S^1$):

$$f : t \mapsto (e^{2\pi i t}, e^{2\pi\sqrt{2}t}).$$

The image of $\mathbb{R}$ under f is dense in $S^1 \times S^1$, and f is an immersion which is not an embedding. $\square$

Definition 3.24. Rank
The rank of a map at a point is defined to be the rank of its differential on the tangent space at that point.

Note that immersions have constant rank (in fact, they have rank equal to the dimension of their domain). Before we move on to the constant rank theorem, we state the following proposition.

Proposition 3.25.
If $f : M \to N$ is smooth and $\operatorname{rank}(f) = k$ at some point $m_0 \in M$, then for all m sufficiently close to m_0, $(rank f_m) \geq k$.

Proof. The rank of f at m_0 is the rank of the matrix $((\frac{\partial f_i}{\partial x_j}(m_0)))$. By a suitable rearrangement of components and coordinates, we may assume that $\det((\frac{\partial f_i}{\partial x_j}(m_0)))_{i,j\leq k} \neq 0$. Since the determinant is a continuous mapping, this deteminant is also non-zero for points sufficiently close to m_0. $\square$

The following theorem, which is a generalization of the Implicit Function Theorem, applies in particular to immersions, but we state the more general version.

Theorem 3.26. Rank Theorem
Suppose that $f : M \to N$ has rank k at all points $m \in M$. Then for all $m \in M$ there are coordinate charts (U, ϕ) containing m and (V, ψ) containing $f(m)$ such that

$$(\psi \circ f \circ \phi^{-1})(r_1, \cdots, r_n) = (r_1, \cdots, r_k, 0, \cdots, 0).$$

Proof. As before, we may assume that $M = \mathbb{R}^m$ and $N = \mathbb{R}^n$. Suppose that $f : M \to N$ has rank k, $f(0) = 0$, and $\det((\frac{\partial f_i}{\partial x_j}(0))_{i,j\leq k}) \neq 0$.
Consider $h : \mathbb{R}^m \to \mathbb{R}^m$ given by $h(x_1, \ldots, x_m) = (f_1(x), \ldots, f_k(x), x_{k+1}, \ldots, x_m))$. Then

$$dh_0 = \begin{pmatrix} (\frac{\partial f_i}{\partial x_j}(0)) & * \\ 0 & I \end{pmatrix}$$

Then h is invertible in a neighborhood U of 0, and $h(0) = 0$. Now let $g = f \circ (h\Big|_U)^{-1}$, so that

$$g(z_1, \ldots, z_n) = (z_1, \ldots, z_k, g_{k+1}(z), \ldots, g_n(z))$$

and $g(0) = 0$. Now, note that

$$dg_z = \begin{pmatrix} I & 0 \\ * & A(z) \end{pmatrix}$$

where $A(z) = (\frac{\partial f_i}{\partial x_j}(0))$, $k+1 \leq i \leq n, k+1 \leq j \leq m$. Furthermore, as the ranks of g and f are both k, the rank of $A(z)$ must be 0, which says that $\frac{\partial g_i}{\partial z_j} = 0$ for $k+1 \leq i \leq n$ and $k+1 \leq j \leq m$. Now consider $K : V \to \mathbb{R}^n$ (where V is some neighborhood of 0) defined by

$$K(y_1, \ldots, y_n) = (y_1, \ldots, y_k, y_{k+1} - g_{k+1}(y_1, \ldots, y_k, 0, \ldots, 0), \ldots, y_n - g_n(y_1, \ldots, y_k, 0, \ldots, 0)).$$

Now, note that

$$dK_0 = \begin{pmatrix} I & 0 \\ * & I \end{pmatrix}$$

which says that K is a diffeomorphism near 0. Finally,

$$\begin{aligned} (K \circ f \circ h^{-1})(z_1, \ldots, z_m) &= K(g(z_1, \ldots, z_n)) \\ &= K(z_1, \ldots, z_k, g_{k+1}(z), \ldots, g_n(z)) \end{aligned}$$

$$= (z_1, \ldots, z_k, g_{k+1}(z) - g_{k+1}(z_1, \ldots, z_k, 0, \ldots, 0), \ldots, g_n(z) - g_n(z_1, \ldots, z_k, 0, \ldots, 0).$$

But since $\frac{\partial g_i}{\partial z_j} = 0$ for $i \geq k+1$, we must have

$$g_i(z) - g_i(z_1, \ldots, z_k, 0, \ldots, 0) = 0 \; i \geq k+1.$$

Combining these statments, we have

$$(K \circ f \circ h^{-1})(z_1, \ldots, z_m) = (z_1, \ldots, z_k, 0, \ldots, 0),$$

whence the desired statement holds. □

Note that since immersions have constant rank, the theorem applies in particular to them.

Exercise 3.2.
Define $f : R^3 \to \mathbb{R}^6$ by $f(x, y, z) = (x^2, y^2, z^2, yz, zx, xy)$. Is f an immersion? Show that the restriction of f to S^2 is an immersion of S^2 into $\mathbb{R}^6$.

Exercise 3.3.
There is no immersion $f : S^2 \to \mathbb{R}^2$.

Exercise 3.4.
(a) Let N be a manifold. Prove that the diagonal $\Delta_N = \{(n, n) \in N \times N : n \in N\}$ is a submanifold of $N \times N$.
(b) Let $F : M \to N$ and $g : L \to N$ be smooth maps such that, for all $m \in M$ and $l \in L$ with $f(m) = g(l)$ we have

$$df_m(T_m M) + dg_l(T_l L) = T_r N, \; r = f(m) = g(l).$$

Show that

$$Z = \{(m, l) \in M \times L : f(m) = g(l)\}$$

is a submanifold of $M \times L$.

Exercise 3.5.
Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a smooth map such that for every x with $||x|| \geq 2$, we have $||f(x)|| < 1/||x||$.
(a) $||f||$ attains its maximum value.
(b) f is not an immersion.

Exercise 3.6.
Let N be a closed submanifold of M. Show that every vector field X on N can be extended to a vector field Y on M.
Hint: First extend the vector field in adapted coordinates. Next, use a partition of unity to combine each of the locally defined extensions into a global vector field.

Exercise 3.7.
Consider $f(x, y) = y^2 + \frac{1}{6}x^6 - \frac{1}{2}x^2$ on $\mathbb{R}^2$. For each $c \in \mathbb{R}$, determine whether or not $f^{-1}(c)$ is a submanifold of $\mathbb{R}^2$. Justify your answer.

4 Vector Fields and Flows

4.1 The Correspondence between Flows and Vector Fields

Definition 4.1. Flow
A **(global) flow** on a manifold M is a map $\Phi : \mathbb{R} \times M \to M$ such that for all $p \in M$ and $s, t \in R$,
(1) $\Phi(0, p) = p$
(2) $\Phi(t, \Phi(s, p)) = \Phi(s + t, p)$.

Given a flow, the second property above tells us that we get a group homomorphism of $\mathbb{R}$ into the group of diffeomorphisms on M by $t \mapsto \Phi_t$; for this reason, the second property is called the **group property** of flows.

Example 4.2.
Let $M = \mathbb{R}$. Then the following are examples of flows:
(1) $\Phi(t, p) = p \cdot e^t$
(2) $\Psi(t, p) = t + p$.

Definition 4.3. Local Flow
A **local flow** is a map $\Phi : A \to M$, where A is open and $A \supset \{0\} \times M$, that satisfies the two flow properties in its domain of definition.

Definition 4.4. Integral Curve
Let X be a vector field on M. A curve $\gamma : I \to M$ for $0 \in I$, an open interval in $\mathbb{R}$, is an **integral curve** of X through $m_0 \in M$ if
(1) $\gamma(0) = m_0$
(2) $\dot{\gamma}(t) = X(\gamma(t))$.

At this point, recall that $\dot{\gamma}(t) = d\gamma_t(\frac{d}{dt})$. One way to think about integral curves is the following: if I place a particle in a vector field X at point p, then the integral curve describes how it will move on the manifold. The vector field merely describes the tangent vectors (or the velocity) of the particle's trajectory at each point. This is one geometric interpretation of integral curves.

Example 4.5.
Let $M = \mathbb{R}$ again, and let $X(t) = t\frac{d}{dt}$. Then the integral curve of X through p is given by $p \cdot e^t$. Note that $\dot{\gamma}_p(t) = p \cdot e^t \frac{d}{dt} = X(\gamma_p(t))$. $\square$

Now, if we knew that integral curves existed and were unique, then given a vector field X, we would define the corresponding flow by

$$\Phi(s, p) = \gamma_p(s).$$

Conversely, given a local flow $\Phi : A \to M$, we can (locally) define a vector field by

$$X(p) = \frac{\partial}{\partial t}\Phi\Big|_{(0,p)},$$

and we can check that $\gamma_p(t) = \Phi(t,p)$ are integral curves of X. This is precisely the correspondence between flows and vector fields.
Now, there are two important issues to realize here. First, the correspondence between flows and vector fields is only *local*, because flows need not exist for all time. This is illustrated by the following example:

Example 4.6.
Let $M = \mathbb{R}$ and take $X(t) = t^2 \frac{d}{dt}$. $\gamma_p(t)$ is an integral curve through p if

$$\dot{\gamma}_p(t) = (\gamma_p(t))^2$$

subject to initial conditions $\gamma_p(0) = t$. If we integrate both sides with perhaps a separation of variables operation as well, we see that if $p \neq 0$, the (unique) solution to this ordinary differential equation is

$$\gamma_p(t) = -\frac{1}{t - \frac{1}{p}}.$$

That is, the flow does not exist for all time unless $p = 0$ (where the solution to the ODE is just $\gamma_0(t) = 0$). $\square$

The other problem with our method above is that we have not established the uniqueness of integral curves. As one might expect, we will do this by way of the existence and uniqueness theorem for ordinary differential equations with smooth dependence on inital conditions. This is part of the content of the next theorem, in which we describe the correspondence between vector fields and flows and check that the equations given above do indeed make sense.

Theorem 4.7.
There is a (local) one-to-one correspondence between vector fields and local flows.

Proof. For the forward direction, let $\Phi(t,m)$ be a local flow on M. Define a vector field X on M by

$$X(p) = \frac{d}{dt}\Big|_{t=0}\Phi(t,p).$$

We claim that $\gamma_p(t) = \Phi(t,p)$ is an integral curve of X; i.e.,

$$\frac{d}{ds}\Big|_{s=t}\gamma_p(s) = X(\gamma_p(t)).$$

To see this, pick a function $f \in C^\infty(M)$. We want to show that

$$\frac{d}{ds}\Big|_{t=s}(f \circ \gamma_p)(s) = (X(\gamma_p(t)))(f).$$

Now, note that

$$\begin{aligned}
\frac{d}{ds}\Big|_{t=s}(f \circ \gamma_p)(s) &= \frac{d}{ds}\Big|_{s=0}(f \circ \gamma_p)(t+s) \\
&= \frac{d}{ds}\Big|_{s=0} f(\Phi(t+s,p)) \\
&= \frac{d}{ds}\Big|_{s=0} f(\Phi(s,\Phi(t,p))) \\
&= [X(\Phi(t,p))](f) = [X(\gamma_p(t))](f).
\end{aligned}$$

Now suppose that X is a vector field on M. We need to show the existence of a local flow, and we want this local flow to be unique. Suppose first that $M \subset \mathbb{R}^n$ is open. Then $TM \simeq M \times \mathbb{R}^n$; let $(x_1, \ldots, x_n)$ denote the global coordinates. Then

$$X(p) = (p, \sum X^i(p)\frac{\partial}{\partial x_i}),$$

where each $X^i \in C^\infty(M)$. $\gamma_p(t) = (\gamma_p^1(t), \ldots, \gamma_p^n(t))$ is an integral curve of X if and only if

$$\frac{d}{ds}\Big|_{s=t}\gamma_p(s) = X(\gamma_p(t)).$$

Using the fact that $\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\}$ forms a basis for the tangent space at each point p, we see that each component γ^i must satisfy

$$(\gamma_p^i)' = X^i(\gamma_p^1(t), \ldots, \gamma_p^n(t))$$

subject to the initial conditions

$$\gamma_p^i(0) = p_i.$$

Now, from the theory of ordinary differential equations, existence and uniqueness and smooth dependence of initial conditions tells us the following:

(1) For any fixed $p \in M$, there is $r > 0$ and $\epsilon > 0$ such that for all $q \in B_r(p)$, there is a curve $\gamma_q : (-\epsilon, \epsilon) \to M$ with

$$\frac{d}{dt}(\gamma_q^i) = X^i(\gamma_q) \;\; \gamma_q(0) = q.$$

(2) The map $\Phi : (-\epsilon, \epsilon) \times B_r(p) \to M$ given by $(t, q) \mapsto \gamma_q(t)$ is smooth
(3) Each $\gamma_q(t)$ is unique.
Thus, integral curves exist and are unique. Now we need to check that Φ is, in fact, a local flow. To do this, first let $\sigma(t) = \Phi(t + s, q)$. We want to show that $\sigma(t) = \gamma_{\Phi(s,q)}(t)$. First, let's check that σ is an integral curve: $\sigma(0) = \Phi(s, q)$, and

$$\begin{aligned}
\frac{d}{dt}\Big|_{t_0} \sigma(t) &= \frac{d}{dt}\Big|_{t_0} \Phi(t + s, q) \\
&= \frac{d}{dt}\Big|_{t_0+s} \Phi(t, q) \\
&= \frac{d}{dt}\Big|_{t_0+s} \gamma_q(t) \\
&= X(\gamma_q(t_0 + s)) \\
&= X(\sigma(t_0)).
\end{aligned}$$

Thus, σ is an integral curve of X through the point $\Phi(s, q)$. By uniqueness, $\sigma(t) = \gamma_{\Phi(s,q)}(t)$. This gives us $\Phi(t, \Phi(s, q)) = \Phi(t + s, q)$. Finally, there is a set $A \subset \mathbb{R} \times B_r(p)$ such that $\Phi : A \to B_r(p)$ is a local flow (take $A = \Phi^{-1}(B_r(p))$, for example, which is open since Φ is continuous.)
This gives us flows locally on $\mathbb{R}^n$. To prove the existence and uniqueness of flows in the general case, first observe that if $\psi : M \to N$ is a diffeomorphism and $X : M \to TM$ is a vector field, we have

$$(d\psi(X))(q) = d\psi_{\psi^{-1}(q)}(X(\psi^{-1}(q))).$$

Note also that if ψ is a flow of X, then $\psi \circ \Phi(\psi \times \iota)$ is a flow of $d\psi(X)$, where ι denotes the identity.
Thus, if (U, ψ) is a coordinate chart with $\psi(U) = B_r(p)$, then if given a vector field X on M, we obtain a flow of $d\psi(X\Big|_U)$, which in turn implies that we get a flow Φ^U of $X\Big|_U$. We need to patch Φ^U into a flow on the manifold (i.e., show that Φ is well-defined), and we'll be done.
To this end, let U and V be two open sets with flows Φ^U and Φ^V, respectively. We want to show that for all $x \in U \cap V$, $\Phi^U(t, x) = \Phi^V(t, x)$, whenever

both sides make sense. Let $\gamma(t) = \Phi^U(t,x)$ and $\sigma(t) = \Phi^V(t,x)$, where $\gamma : I_1 \to M$ and $\sigma : I_2 \to M$. Let $J = \{t \in I_1 \cap I_2 : \gamma(t) = \sigma(t)\}$. Note that J is open because of uniqueness of solutions of ODEs, and J is closed since $J = (\gamma \times \sigma)^{-1}(\Delta_M)$ is a closed set. Thus, since $I_1 \cap I_2$ is connected, $J = I_1 \cap I_2$. $\square$

To summarize, the correspondence between vector fields X near a point p and integral curves γ through p is given by following system of ODE's :

$$X^i(\gamma_p) = \dot{\gamma}_p^i$$

with initial conditions

$$\gamma_p^i(0) = p^i.$$

The existence and uniqueness theorem for ordinary differential equations guarantees that local solutions exist and are unique. Thus, insomuch as we can solve ODE's, we can solve this system to obtain a general form for integral curves through a point p in order to define a local flow.
Also, note that if $\Phi_1 : A_1 \to M$ and $\Phi_2 : A_2 \to M$ are local flows of X, then so is their "union" $\Phi : A_1 \cup A_2 \to M$. (We essentially showed this in the last proof). Therefore, it makes sense to speak of the **maximal flow** of X on the manifold, which is defined to be the union of all flows.

Exercise 4.1.
Find the flows of the following vector fields on $\mathbb{R}^2$:
(1) $X = x_1\frac{\partial}{\partial x_1} + x_2\frac{\partial}{\partial x_2}$.
(2) $Y = x_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1}$.

Exercise 4.2.
Prove that if a vector field X on a manifold M satisfies $X(m_0) = 0$ for some $m_0 \in M$, then there is an open set W containing m_0 such that the flow of X on W exists for all $t \in [0,1]$.

Exercise 4.3.
Consider the following vector fields on $\mathbb{R}^4 - \{0\}$.
(1) $X_1 = -x_2\frac{\partial}{\partial x_1} + x_1\frac{\partial}{\partial x_2} + x_4\frac{\partial}{\partial x_2} - x_3\frac{\partial}{\partial x_4}$
(2) $X_2 = -x_3\frac{\partial}{\partial x_1} - x_4\frac{\partial}{\partial x_2} + x_1\frac{\partial}{\partial x_2} + x_2\frac{\partial}{\partial x_4}$
(3) $X_3 = -x_4\frac{\partial}{\partial x_1} + x_3\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_2} + x_1\frac{\partial}{\partial x_4}$
Show the following:
(a) Each X_i defines a smooth, non-vanishing vector field on S^3.
(b) For each $p \in S^3$, $\{X_i(p)\}$ forms a basis for T_pS^3.
(c) Compute the flows for X_1.

(d) Let $f : S^3 \to \mathbb{R}$ be given by $f(x_1, x_2, x_3, x_4) = x_3x_1 - x_2x_4$; show that f is constant along integral curves of X_1.

Exercise 4.4.
Let M be a manifold. An **isotopy** on M is a collection of diffeomorphisms $\{f_t : M \to M\}_{t\in(-\epsilon,\epsilon)}$ such that
1) f_0 is the identity.
2) The map $(-\epsilon, \epsilon) \times M \to M$ given by $(t, m) \mapsto f_t(m)$ is smooth.
A **time-dependent vector field** $\{X_t\}$ is a smooth map $(-\epsilon, \epsilon)\times M \to TM$ given by $(t, m) \mapsto X_t(m)$. An isotopy defines a time-dependent vector field by

$$X_s(f_s(m)) = \frac{d}{dt}\Big|_{t=s} f_t(m).$$

Prove: given a time-dependent vector field $\{X_t\}$, there is an isotopy $\{f_t\}$ such that the previous equation holds.
Hint: Let $\overline{X}(t, m) = (\frac{d}{dt}, X_t(m))$, a vector field on $\mathbb{R} \times M$. Its local flow $\Phi_s(t, m)$ is of the form $\Phi_s(t, m) = (\Phi^1_s(t, m), \Phi^2_s(t, m))$. Show that $\Phi^2_s(t, m) = s + t$.

Exercise 4.5.
Consider a time-dependent vector field $X_t(m) = t\frac{d}{d\theta}$ on S^1. Compute the corresponding isotopy.

The fact that global flows need not exist motivates the following definition.

Definition 4.8. Complete Vector Field
A vector field is **complete** if the maximal flow exists for all time.

Example 4.9.
We have already shown that the vector field $X(p) = p^2\frac{d}{dt}$ has a flow which does not exist for all time. Hence, X is not complete. □

The next theorem asserts that any vector field with compact support is complete. Before we prove this, we need the following lemma.

Lemma 4.10.
Let X be a vector field on M with flow $\Phi(t, x)$. Suppose that there is an element $a \in \mathbb{R}$ such that $\Phi(a, x)$ exists for all $x \in M$. Then $(d\Phi_a)_x(X(x)) = X(\Phi_a(x))$, where $\Phi_a(x) = \Phi(a, x)$.

Proof. Fix $f \in C^\infty(M)$. Then

$$((d\Phi_a)_x)(X(x))(f) = (X(x))(f \circ \Phi_a).$$

Furthermore, for all $y \in M$, we have

$$(X(y))(f) = \frac{d}{dt}\Big|_{t=0} f(\Phi(t, y)),$$

by the definition of a flow. Therefore,

$$\begin{aligned}
(X(x))(f \circ \Phi_a) &= \frac{d}{dt}\Big|_{t=0} (f \circ \Phi_a)(\Phi(t, x)) \\
&= \frac{d}{dt}\Big|_{t=0} f(\Phi(a, \Phi(t, x))) \\
&= \frac{d}{dt}\Big|_{t=0} f(\Phi(a + t, x)) \\
&= \frac{d}{dt}\Big|_{t=0} f(\Phi(t, \Phi_a(x))) \\
&= X(\Phi_a(x))(f).
\end{aligned}$$

□

Theorem 4.11.
Let X be a vector field on a manifold M. If its support is compact, then it is a complete vector field.

Proof. Note first that if $X(x) = 0$, then $\Phi(t, x) = 0$ for all t. So on $M - \mathrm{supp}(X)$, the flow of X is defined for all t. Let $A \subset \mathbb{R} \times M$ be the domain of the flow Φ of X. As $\{0\} \times \mathrm{supp}(X) \subset A$ is compact and A is open, there is a number $\epsilon > 0$ such that $(-\epsilon, \epsilon) \times \mathrm{supp}(X) \subset A$. If Φ did not exist on $[\epsilon, 2\epsilon]$, then we could define it by $\Phi_1(t, x) = \Phi(\epsilon, \Phi(t - \epsilon, x)) = \Phi_\epsilon(\Phi(t - \epsilon, x))$. We claim that $\Phi_1(t, x)$ is a flow on X.

$$\begin{aligned}
\frac{d}{dt}\Big|_{t=0} \Phi_1(t, x) &= \frac{d}{dt}\Big|_{t=0} \Phi_\epsilon(\Phi(t - \epsilon, x)), \quad t \in [\epsilon, 2\epsilon] \\
&= (d\Phi_\epsilon)(X(\Phi(t_0 - \epsilon, x))) \\
&= X(\Phi_\epsilon(\Phi(t_0 - \epsilon, x))) \\
&= X(\Phi(\epsilon + (t_0 - \epsilon), x)) \\
&= X(\Phi_1(t_0, x)).
\end{aligned}$$

Hence, $\Phi_1 = \Phi\Big|_{[\epsilon, 2\epsilon] \times M}$, by uniqueness of flows and the maximality of Φ_1. By induction, we can then see that $[k\epsilon, (k+1)\epsilon] \times M \subset A$ for all $k \in \mathbb{N}$. □

4.2 Lie Derivatives

Now we discuss one way to differentiate vector fields.

Definition 4.12. Lie derivative
Let X and Y be vector fields on a manifold M. Let ϕ_t denote the (local) flow of X. The **Lie derivative** $L_X Y$ of Y with respect to X is a vector field on M given by

$$(L_X Y)(p) = \lim_{t \to 0} \frac{1}{t}((d\Phi_{-t})_x(Y(\Phi_t(p)) - Y(p))) = \frac{d}{dt}\Big|_{t=0} (d\Phi_{-t})_x(Y(\Phi_t(p))).$$

We can think of the Lie derivative as differentiating the vector field Y with respect to the vector field X. As it turns out, this is an already familiar object, the Lie bracket.

Proposition 4.13.
$L_X Y = [X, Y]$.

Proof. Observe that if $\gamma : I \to T_pM$ is a curve, then for any $f \in C^\infty(M)$,

$$(\frac{d}{dt}\Big|_{t=0}\gamma(t))(f) = \frac{d}{dt}\Big|_{t=0}(\gamma(t)f).$$

To see this, simply compute both sides in coordinates, and you'll get the same answer on both sides.
Now, fix $f \in C^\infty(M)$.

$$\begin{aligned} ((L_X Y)(p))(f) &= \frac{d}{dt}\Big|_{t=0}((d\Phi_t)(Y(\Phi_t(p)))(f)) \\ &= \frac{d}{dt}\Big|_{t=0}(Y(\Phi_t(p))(f \circ \Phi_{-t})). \end{aligned}$$

Now denote the flow of Y by Ψ_s. Then for all $q \in M$,

$$(Y(q))(f) = \frac{d}{ds}\Big|_{s=0} f(\Psi_s(q)).$$

Thus,

$$\begin{aligned} ((L_X Y)(p))(f) &= \frac{d}{dt}\Big|_{t=0}(\frac{\partial}{\partial s}\Big|_{s=0}(f \circ \Phi_{-t})(\Psi_s(\Phi_t(p)))) \\ &= \frac{\partial^2}{\partial t \partial s}\Big|_{(0,0)}(f \circ \Phi_{-t} \circ \Psi_s \circ \Phi_t)(p). \end{aligned}$$

Let $H = (f \circ \Phi_{-t} \circ \Psi_s \circ \Phi_t)(p)$. Let $G(s,a,b) = (f \circ \Phi_a \circ \Psi_s \circ \Phi_b)$. Then $H(s,t) = G(s,-t,t)$. Also,

$$\begin{aligned}
\frac{\partial H}{\partial s}(s,t) &= \frac{\partial G}{\partial s}(s,-t,t) \\
&= \frac{\partial^2 H}{\partial t \partial s}(s,t) \\
&= -\frac{\partial^2 G}{\partial a \partial s}(0,0,0) + \frac{\partial^2 G}{\partial b \partial s}(0,0,0).
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\frac{\partial^2 G}{\partial a \partial s}(0,0,0) &= \frac{\partial^2 G}{\partial s \partial a}(0,0,0) \\
&= \frac{\partial}{\partial a}\Big|_{s=0}\left(\frac{\partial G}{\partial a}\right)(s,0,0) \\
&= \frac{\partial}{\partial s}\Big|_{s=0}\frac{\partial}{\partial a}\Big|_{a=0}(f \circ \Phi_a)(\Psi_s(p)) \\
&= \frac{\partial}{\partial s}\Big|_{s=0}(X(f))(\Psi_s(p)) \\
&= (Y(X(f)))(p).
\end{aligned}$$

Thus, we have that

$$-\frac{\partial^2 G}{\partial a \partial s}(0,0,0) = -Y(X(f)).$$

Similarly, we get that

$$\frac{\partial^2 G}{\partial b \partial s}(0,0,0) = X(Y(f)).$$

Thus, it follows that $L_X Y = [X,Y]$. □

The following proposition is a nice geometric characterization of what it means for the Lie bracket to be zero.

Proposition 4.14.
Let Φ_t and Ψ_s denote local flows of X and Y, respectively. Then

$$[X,Y] = 0 \textit{ if and only if } \Phi_t \circ \Psi_s = \Psi_s \circ \Phi_t.$$

Proof. First suppose that $\Phi_t \circ \Psi_s = \Psi_s \circ \Phi_t$ whenever both sides make sense. Now, for all $x \in M$,

$$\frac{d}{dt}\Big|_{t=0} \Phi_t(\Psi_s(s)) = \frac{d}{dt}\Big|_{t=0} (\Psi_s \circ \Phi_t)(x).$$

Thus,

$$X(\Psi_s(x)) = (d\Psi)(X(x)).$$

In turn, we see that

$$d\Psi_{-s}(X(\Psi_s(x))) = X(x)$$

for all s, which means that

$$0 = \frac{d}{ds}\Big|_{s=0} (d\Psi_{-s})(X(\Psi_s(x))) = L_Y X = [Y, X] = -[X, Y].$$

For the converse, suppose that $[X, Y] = 0$. Then

$$0 = \frac{d}{dh}\Big|_{h=0} (d\Phi_{-h})(Y(\Phi_h(p)))$$

for all $p \in M$, which implies that

$$\begin{aligned}
\frac{d}{dt}\Big|_{t=s} (d\Phi_{-t})(Y(\Phi_{-t}(p))) &= \frac{d}{dh}\Big|_{h=0} (d\Phi_{-s+h})(Y(\Phi_{s+h}(p))) \\
&= \frac{d}{dh}\Big|_{h=0} (d\Phi_{-s})[(d\Phi_{-h})(Y(\Phi_h(\Phi_s(p))))] \\
&= (d\Phi_{-s})[\frac{d}{dh}\Big|_{h=0} (d\Phi_{-h})(Y(\Phi_h(\Phi_s(p))))] \\
&= 0.
\end{aligned}$$

(Here, we use the fact that $d\Phi$ is linear, viewing $Y(\Phi_h(\Phi_s(p))$ as a curve in $T_{\Phi_s(p)}M$. As an aside, note that if $A : V \to W$ is linear with V,W finite-dimensional, then $\frac{d}{dt}(A(\gamma(t))) = A(\frac{d}{dt}\gamma(t))$.) So

$$d\Phi_{-t}(Y(\Phi_t(p))) = d\Phi_0(Y(\Phi_0(p)))$$

for sufficiently small t. But

$$d\Phi_0(Y(\Phi_0(p))) = Y(p)$$

so that

$$Y(\Phi_t(p)) = (d\Phi_t(p))(Y(p));$$

i.e., Y is constant along integral curves of X.
Now, fix $p \in M$ and $t \in \mathbb{R}$. Consider $\gamma(s) = \Phi_t(\Psi_s(p))$. As $\gamma(0) = \Phi_t(p)$, we can then compute the derivative of γ with respect to s to show that γ is an integral curve of Y and then apply the uniqueness of integral curves. Explicitly,

$$\begin{aligned} \frac{d}{ds}\Big|_{s=s_0}\gamma &= \frac{d}{ds}\Big|_{s=s_0}[\Phi_t(\Psi_s(p))] \\ &= (d\Phi_t)(\frac{d}{ds}\Big|_{s=s_0}\Psi_s(p) \\ &= (d\Phi_t)(Y(\Psi_s(p))) = Y((\Phi_t \circ \Psi_s)(p)) \\ &= Y(\gamma(s_0)). \end{aligned}$$

So $\gamma(s)$ is an integral curve through $\Phi_t(p)$. By uniqueness of integral curves,

$$\Phi_t(\Psi_s(p)) = \gamma(s) = \Psi_s(\Phi_t(p)).$$

□

Exercise 4.6.
Suppose that M and N are manifolds. Show that if $X \in \Gamma(TM)$ and $Y \in \Gamma(TN)$, then $[X, Y] = 0$ on $M \times N$.

Exercise 4.7.
Suppose that X and Y are vector fields on M. Compute an expression for $[X, Y]$ in local coordinates.

4.3 F-related Vector Fields

Consider the following example:

Example 4.15. Let $M = \mathbb{R}^2$, and define $X(x, y) = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}$. This is the circular flow around the origin. Now consider the natural projection $F : \mathbb{R}^2 \to \mathbb{R}$. Note that $dF_x(X(x))$ is not a well-defined vector field on $\mathbb{R}$; that is, we cannot use F to push X forward. □

Now suppose that $F : M \to M'$ is a diffeomorphism, and $X \in \Gamma(TM)$, $X' \in \Gamma(TM')$. Then $X'(m') = (dF)_{F^{-1}(m)}(X(F^{-1}(m')))$ is a vector field on M; that is, the following diagram commutes:

$$\begin{tikzcd} TM \arrow[r, "dF"] & TM' \\ M \arrow[u, "X"] \arrow[r, "F"] & M' \arrow[u, "X'"'] \end{tikzcd}$$

That is, we can push vector fields forward by diffeomorphisms but not by arbitrary smooth maps, as the following example shows. This leads to the concept of two vector fields being **F-related**.

Definition 4.16. F-related
Let $F : M \to N$ be a map, and let X and Y be vector fields on X and Y, respectively. X and Y are **F-related** if $dF \circ X = Y \circ F$.

Note that if $F : M \to N$ is a diffeomorphism and X is any vector field on M, then X is F-related to a *unique* vector field on N. However, in general, one vector field on M can be F-related to many different vector fields on N.

Example 4.17.
Vector fields on a submanifold are F-related to those on the larger manifold if and only if F is the inclusion. □

To conclude the section, we state the following lemma, which will be of some use in certain situations later on in the course.

Lemma 4.18.
Let $F : M \to N$ be a map, and suppose that X and Y are F-related to W and Z, respectively. Then $[X, Y]$ is F-related to $[W, Z]$.

Proof. For any $h \in C^\infty(N)$,

$$\begin{aligned}(Y(h \circ F))(x) &= Y(x)(h \circ F) \\ &= dF_x(Y(x))(h) \\ &= Z(F(x))(h) \\ &= Z(h)(F(x)) \\ &= (Z(h) \circ F)(x).\end{aligned}$$

That is, if Y and Z are F-related, then for all $h \in C^\infty(N)$, $Y(h \circ F) = Z(h) \circ F$. Similarly, $X(h \circ F) = W(h) \circ F$ for all $h \in C^\infty(N)$. Thus,

$$\begin{aligned}([X, Y])(h \circ F) &= X(Y(h \circ F)) - Y(X(h \circ F)) \\ &= X(Z(h) \circ F) - Y(W(h) \circ F) \\ &= W(Z(h)) \circ F - Z(W(h)) \circ F \\ &= ([W, Z](h) \circ F).\end{aligned}$$

Thus, we have that

$$\begin{aligned}[(dF_x)([X,Y])(x)](h) &= ([X,Y])(h \circ F)(x) \\ &= ([W,Z](h) \circ F)(x) \\ &= ([W,Z](F(x)))(h).\end{aligned}$$

Therefore,

$$(dF_x)[X,Y] = [W,Z] \circ F.$$

□

Exercise 4.8.
Let $F : M \to N$ be a smooth map of manifolds, and let X and Y be vector fields on M and N, respectively, which are F-related. Show that any integral curve of X is mapped by F to an integral curve of Y.

Exercise 4.9.
Suppose that N is a submanifold of M. Let Z and Y be two vector fields on M such that for all $x \in N$, $Z(x), Y(x) \in T_xN$. Show that $[Z,Y](x) \in T_xN$ as well.

5 Vector Bundles

5.1 Basic Definitions

Definition 5.1. Vector Bundle
A **real vector bundle** E over a manifold M of rank k is a disjoint union of smoothly varying k-dimensional vector spaces E_x, $x \in M$. Specifically, if E is a vector bundle over M, then E has the following ingredients :
(1) E is a manifold
(2) The projection $\pi : E \to M$ given by $\pi[E_x] = \{x\}$ is smooth
(3) E is **locally trivial**: For all $x \in M$, there is an open neighborhood U_x of $x \in M$ and a diffeomorphism $\psi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ such that $pr \circ \psi = \pi$, where pr is the natural projection from $U \times \mathbb{R}^k$ to U. In addition, we have that $\psi\Big|_{E_y} : E_y \to \{y\} \times \mathbb{R}^k$ is an isomorphism for all $y \in U$.
E is called the **total space**.
M is called the **base space**.
E_x is called a **fiber** of $\pi : E \to M$.
The maps $\psi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ are called **local trivializations.**

Thus, a vector bundle is a manifold that is locally trivial; i.e., locally, it is a projection. However, it may not be a *global* projection, for it may somehow be "twisted". The nature of this twisting is very important to us. We can similarly define complex vector bundles, by the way, and every complex vector bundle would be a real vector bundle of rank $2k$.

Example 5.2.
The projection $\pi : M \times \mathbb{R}^k \to M$ is a vector bundle of rank k. It is an example of a **trivial vector bundle**. □

Definition 5.3.
A vector bundle over M is a **trivial vector bundle** if it is isomorphic to $M \times \mathbb{R}^k$, where $k = \text{rank}(E)$.

Example 5.4.
Let's look at the tangent bundle. We already know that TM is a manifold and $\pi : TM \to M$ is smooth. What are trivializations? For each $x \in M$, take a coordinate chart $(U, \phi = (x_1, \ldots, x_n))$. Then define $\psi : \pi^{-1}(U) \to U \times \mathbb{R}^n$ by

$$(q, v) \mapsto (q, (dx_1)(v), \ldots, (dx_n)(v)),$$

where $v \in T_qM$, of course. Note that

$$\psi^{-1}(q, v_1, \ldots, v_n) = (q, \sum_i v_i \frac{\partial}{\partial x_i}).$$

$\square$

Example 5.5.
TS^1 is trivial : $(\theta, x) \mapsto (\theta, x\frac{d}{d\theta})$ is a global trivialization. (Alternatively, $x_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1}$ is a vector field on $\mathbb{R}^2$ whose restriction to S^1 is a nowhere-vanishing vector field on S^1.) $\square$

Example 5.6.
As earlier mentioned, $\pi : T^*M \to M$ is a vector bundle. We'll consider this example in greater detail later on. $\square$

Example 5.7.
Recall $\mathbb{C}P^n$ is the set of all complex lines in $\mathbb{C}^{n+1}$, which we identify with $\{\mathbb{C}^{n+1} - \{0\}\}/$, where two lines l, l' are said to be equivalent if there is a complex number $\lambda \neq 0$ such that $l = \lambda l'$. Define

$$L = \{(l, v) \in \mathbb{C}P^n \times \mathbb{C}^{n+1} : v \in l\}.$$

$\pi : L \to \mathbb{C}P^n$ is called the **tautological complex line bundle**. It is a complex vector bundle of rank 1, and hence a real vector bundle of rank 2. To see this, note that $v \in [w]$ if and only if there is a $\lambda \in C^x$ such that $\lambda v = w$. So if $v = (v_1, \ldots, v_{n+1})$, $w = (w_1, \ldots, w_{n+1})$ with $w = \lambda v$, then $\lambda = w_i/v_i = w_j/v_j$ for all i, j. That is, we have $v_j w_i = v_i w_j$ for all i, j. Thus, $L = \{([w], v) : v_j w_i = v_i w_j\}$. From this, one can get that L is a manifold.
As for trivializations, let $U_i = \{[w] \in \mathbb{C}P^n : w_i \neq 0\}$. Define $\Psi_i : \pi^{-1}(U_i) \to U_i \times \mathbb{C}$ by

$$\Psi_i([w], v) = ([w], v_i).$$

Then

$$\psi_i^{-1}([w], z) = ([w], z(\frac{w_1}{w_i}, \ldots, \frac{w_{n+1}}{w_i})).$$

$\square$

Here is a the appropriate notion of a vector bundle map. Not only should it be a smooth map of manifolds, but it should also be an isomorphism when restricted to fibers.

Definition 5.8. Bundle Map
Let $\pi_E : E \to M$ and $\pi_F : F \to M$ be two vector bundles over a manifold M. A smooth map $f : E \to F$ is a **map of vector bundles** if $f(E_x) \subset F_x$ for all x and if $f\Big|_{E_x} : E_x \to F_x$ is linear.

Definition 5.9. If f is a map of vector bundles and f^{-1} exists, then f is a **bundle isomorphism.**

Note that we do not require $g^{-1}\Big|_{F(x)}$ to be linear, because it's automatic. That is, if $f : M \times \mathbb{R}^k \to M \times \mathbb{R}^k$ is given by

$$(m, v) \mapsto (m, A(m)v)$$

where $A : M \to \mathrm{GL}(\mathbb{R}, k)$, then

$$f^{-1}(m, w) = (m, ((A(m))^{-1}w).$$

Exercise 5.1.
Let f be a vector bundle map that is an isomorphism on each fiber. Prove that f^{-1} is smooth.
Hint: Prove that $\mathrm{inv} : \mathrm{GL}(\mathbb{R}, k) \to \mathrm{GL}(\mathbb{R}, k)$ given by $A \mapsto A^{-1}$ is smooth.

Definition 5.10. Section
A **section** $s : M \to E$ of a vector bundle $\pi : E \to M$ is a C^∞ map such that $\pi \circ s = id\Big|_M$. That is, a smooth map from M to E such that $s(x) \in E_x$ for all x.

The collection of sections of E is denoted by $\Gamma(E)$.

Example 5.11.
$s(x) = 0_x \in E_x$ is a section called the zero section. □

Example 5.12.
Vector fields on M are sections of the tangent bundle, hence the notation $\Gamma(TM)$. □

Note that if $s \in \Gamma(E)$ and $f \in C^\infty(M)$, then $(fs)(x) = f(x)s(x)$ is also a section; thus, we see that $\Gamma(E)$ is a $C^\infty(M) - \mathit{module}$.

Definition 5.13. Local Section
A **local section** of $\pi : E \to M$ is a section of $\pi^{-1}[U] = E\Big|_U$ for some open $U \subset M$. Alternatively, a local section is a map $s : U \to E$ such that $\pi \circ s = id_M$.

Example 5.14.
If $(U, x_1, \dots, x_n)$ is a coordinate chart on M, then $q \mapsto \left.\frac{\partial}{\partial x_i}\right|_q$ is a local section of TM. $\square$

Proposition 5.15.
A vector bundle of rank k is trivial if and only if there are sections $s_1, \dots, s_k \in \Gamma(E)$ such that, for all $x \in M$, $\{s_1(x), \cdots, s_k(x)\}$ is a basis of E_x. Such a collection of sections is also called a **frame.**

Proof. Suppose that E is a trivial vector bundle over a manifold M. Then we have a global trivialization $\psi : \pi^{-1}(M) \to M \times \mathbb{R}^k$. Define

$$s_i(x) = \psi_x^{-1}(e_i)$$

where $\{e_1, \dots, e_k\}$ is the canonical basis for $\mathbb{R}^k$. Then the collection $\{s_1, \dots, s_k\}$ satisfies the desired properties.
Conversely, suppose that we have smooth sections $s_1, \dots, s_k$ that form a basis of E_x at every x. Then a global trivialization is given by

$$(q, v_1, \dots, v_k) \mapsto \sum_i v_i s_i(q).$$

$\square$

Exercise 5.2.
Let $x \in M$ and $v \in E_x$. Then there is a *global section* s such that $s(x) = v$.

5.2 Vector Bundles via Transition Maps

Suppose that $\pi : E \to M$ is a vector bundle of rank k, $\{U_\alpha\}$ is a cover of M such that $\psi_\alpha : \pi^{-1}[U_\alpha] \to U_\alpha \times \mathbb{R}^k$ is a local trivialization. If $U_\alpha \cap U_\beta \neq \emptyset$, we have a map

$$\psi_\beta \circ \psi_\alpha^{-1} : (U_\alpha \cap U_\beta) \times \mathbb{R}^k \to (U_\alpha \cap U_\beta) \times \mathbb{R}^k$$

given by

$$(q, v) \mapsto (q, \psi_{\alpha\beta}(q)v).$$

Note that $\psi_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(\mathbb{R}^k)$ is smooth, because for every basis vector e_j of $\mathbb{R}^k$, the map

$$q \mapsto \psi_{\alpha\beta}(q)e_j$$

is smooth. Such maps are called a **transition maps** for the bundle $\pi : E \to M$.

Transition maps characterize the "twisting" in a vector bundle over a manifold; that is, they tell us how the bundle changes over the manifold. Furthermore, it turns out that vector bundles are characterized (up to isomorphism) by their transition maps. We now state this result, but first, we need a couple of definitions.

Definition 5.16. Cocycle Conditions
Consider the maps $\phi_{\alpha\beta}$ from the previous definition. If, for all α, β, and γ we have in $GL(\mathbb{R}^k)$,
(1) $\phi_{\alpha\alpha} = id$
(2) $\phi_{\alpha\beta} \cdot \phi_{\beta\alpha} = id$
(3) $\phi_{\alpha\beta} \cdot \phi_{\beta\gamma} = \phi_{\alpha\gamma}$
the transition maps $\{\phi_{\alpha\beta}\}$ are said to satisfy the **cocycle conditions.**

Theorem 5.17. *Let M be a manifold, $\{U_\alpha\}$ an open cover, and $\{\phi_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(\mathbb{R}^k)\}$ a collection of smooth maps satisfying the cocycle conditions. Then there is a unique vector bundle E over M of rank k with transition maps $\{\phi_{\alpha\beta}\}$.*

Proof. The following is only a sketch of the proof, but from this sketch, one can get at least some idea as to why this theorem is true. Let

$$\overline{E} = \coprod_\alpha (U_\alpha \times \mathbb{R}^k).$$

Define a relation on $\overline{E}$ by

$$(q, v) \sim (q', v') \text{ if and only if } q = q', \psi_{\beta\alpha}(v) = v'.$$

The cocyle conditions hold if and only if this relation is an equivalence relation. Let $E = \overline{E}/\sim$, and write $[q, v]$ for the equivalence class of (q, v). We have $\pi : E \to M$ given by $\pi([q, v]) = q$. Note that $\pi^{-1}(U_\alpha) = \{[q, v] : (q, v) \in U_\alpha \times \mathbb{R}^k\}$, so that we obtain trivializations $\pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^k$ given by $[q, v] \mapsto (q, v)$. Finally, one needs to check that this assignment is well-defined. □

Let's illustrate such a construction via an example.

Example 5.18.
Let $\pi_E : E \to M$ and $\pi_F : F \to M$ be vector bundles. We want to construct

$E \oplus F \to M$ so that $(E \oplus F)_x = E_x \oplus F_x$. To proceed, pick an open cover $\{U_\alpha\}$ of M such that $E\Big|U_\alpha$ and $F\Big|_{U_\alpha}$ are trivial. Let $\psi^E_{\alpha\beta} : U\alpha \cap U_\beta \to \mathrm{GL}(\mathbb{R}^k)$ and $\psi^F_{\alpha\beta} : U\alpha \cap U_\beta \to \mathrm{GL}(\mathbb{R}^l)$ be transition maps. Define

$$\psi^{E \oplus F}_{\alpha\beta}(a) = \begin{pmatrix} \psi^E_{\alpha\beta}(a) & 0 \\ 0 & \psi^F_{\alpha\beta}(a) \end{pmatrix} \in \mathrm{GL}(\mathbb{R}^k \oplus \mathbb{R}^l).$$

$\psi^{E \oplus F}_{\alpha\beta}$ are smooth and satisfy cocycle conditions, as one can check; therefore, there is a vector bundle with transition maps $\psi^{E \oplus F}_{\alpha\beta}$, which we denote by $E \oplus F$. $\square$

Let V and W be finite-dimensional vector spaces, $A \in \mathrm{GL}(V)$ and $B \in \mathrm{GL}(W)$ over M. Then
(1) $A \oplus B \in \mathrm{GL}(V \oplus W)$
(2) $A \otimes B \in \mathrm{GL}(V \otimes W)$
(3) $Hom(E, F) \in \mathrm{GL}(\mathrm{Hom}(V, W))$
(Here, $\mathrm{Hom}(A, B)T = B \circ T \circ A^{-1}$.)
(4) A^*
(Here, $A^*(l) = l \circ A^{-1}$.) (5) $\Lambda^k(A) : \Lambda^k(V) \to \Lambda^k(V)$.

Mirroring the construction given above, we can construct transition maps for the following vector bundles (here, E and F are vector bundles over M:
(1) $E \otimes F$
(2) E^*
(3) $\mathrm{Hom}(E, F)$
(4) $\Lambda^k(E)$.

Exercise 5.3.
Compute transition maps for the tautological real line bundle $L \mapsto \mathbb{R}P^n$:

$$L = \{(l, v) :\in \mathbb{R}P^n \times \mathbb{R}^{n+1} : v \in l\}.$$

Also compute transition maps for $L \otimes L$. (Hint: write down the isomorphism $\mathbb{R} \otimes \mathbb{R} \to \mathbb{R}$.)

Exercise 5.4.
Let $\pi_E : E \to M$ and $\pi_F : F \to M$ be vector bundles over M.
(a) Show that $E \times F$ is a vector bundle over $M \times M$.
(b) Explain why $G = \{(e, f) \in E \times F : \pi_E(e) = \pi_F(f)\}$ can be considered a vector bundle over M.
(c) Show that, as a vector bundle over M, G is isomorphic to $E \oplus F$.

Exercise 5.5.
Let $\pi_E : E \to M$ and $\pi_F : F \to M$ be two vector bundles. Let $TL\Gamma(E) \to \Gamma(F)$ be an $\mathbb{R}$-linear map such that for any $f \in C^\infty(M)$ and any $s \in \Gamma(E)$,

$$T(fs) = fT(s).$$

Show that there is a vector bundle map $\psi : E \to F$ such that $\psi \circ s = T(s)$ for any $s \in \Gamma(E)$.
Hint: given $v \in E_x$, use a previous exercise to find $s \in \Gamma(E)$ with $s(x) = v$. Define $\psi(v) = [T(s)](x)$. Show that ψ is well-defined. You may wish to use local trivializations to check the smoothness of ψ.

5.3 The Cotangent Bundle as a Vector Bundle

Now we return to the cotangent bundle. Our goal here will be to compute transition maps and to show that $(T^*M) = (TM)^*$ (i.e., the cotangent bundle is the dual bundle of the tangent bundle).
First, let's compute transition maps for the tangent bundle. Let U be an open subset of M for which we have coordinates $(x^1, \ldots, x^n)$ and $(y^1, \cdots, y^n)$. Then $\{\frac{\partial}{\partial x_i}\}$ and $\{\frac{\partial}{\partial y_i}\}$ define trivializations of the tangent bundle TU. In particular, we send

$$U \times \mathbb{R}^k \to TU \text{ by } (q, (v_1, \ldots, v_n)) \mapsto \sum v_i \frac{\partial}{\partial x_i}\Big|_q.$$

(Since we can always invert, the direction in which we write this function does not matter.) Now let us compose, so that transition maps are given by

$$(q, (v_1, \ldots, v_n)) \mapsto (q, \ldots, \sum_j v_j \frac{\partial y_i}{\partial x_j}(q), \ldots).$$

That is, the transition maps $\psi : U \to GL(\mathbb{R}^n)$ are given by (where i indexes rows and j indexes columns)

$$q \mapsto (\frac{\partial y_i}{\partial x_j}(q)).$$

Now, on this open set U, $\{dx_i\}$ and $\{dy_j\}$ are both frames of T^*U. We have the following maps :

$$U \times \mathbb{R}^n \to T^*U \text{ by } (q, \eta_1, \ldots, \eta_n) \mapsto (q, \sum \eta_i (dx_i)_q)$$

and

$$T^*U \mapsto U \times \mathbb{R}^n \text{ by } (q, \eta) \mapsto (q, (\eta(\frac{\partial}{\partial y_1}), \ldots, \eta(\frac{\partial}{\partial y_n}))).$$

So the transition maps for the contangent bundle are given by

$$(q, (\eta_1, \ldots, \eta_n)) \mapsto (q, \ldots, \sum \eta_j (dx_j)_q (\frac{\partial}{\partial y_i}), \ldots).$$

That is, our map $U \to GL(\mathbb{R}^n)$ here is given by

$$q \mapsto (\frac{\partial x_j}{\partial y_i}(q)).$$

By the chain rule,

$$\sum \frac{\partial y_i}{\partial x_j} \frac{\partial x_j}{\partial y_k} = \delta_{ik};$$

that is, the matrix

$$(\frac{\partial y_i}{\partial x_j})^{-1} = (\frac{\partial x_i}{\partial y_j}).$$

Thus, the transition maps for the contangent bundle are the inverse transpose of the transition maps for the tangent bundle. This shows that $T^*U \simeq (TU)^*$ as vector bundles.
Here is a similar exercise:

Exercise 5.6.
Let M and N be two manifolds. Show that $T^*(M \times N) \simeq T^*M \times T^*N$.

6 Differential Forms

6.1 The Exterior Derivative

We now change modes a bit and study differential forms. Roughly speaking, a differential form at a point q is an alternating multi-linear map on the tangent space to the manifold at that point. Here's an "official" definition:

Definition 6.1. Differential Form
A **differential form** on a manifold M is an element

$$\omega \in \Gamma(\Lambda^k(T^*M))$$

for some k. Using multi-index notation, locally (i.e., within a coordinate chart), we can write $\omega = \sum_{\left|I\right|=k} a_I dx_I$, where $a_I \in C^\infty(M)$.
By $\Omega^k(M)$, we mean the collection of all k-forms on M, and by $\Omega(M)$, we mean the collection of forms on M.
In particular, $\Omega^0(M) = C^\infty(M)$, the collection of smooth functions from $M \to \mathbb{R}$.

Example 6.2.
Suppose that $M = \mathbb{R}^n$, where $x_1 \ldots, x_n$ are the standard (global) coordinates. Then $dx_1, \ldots, dx_n$ are 1-forms on M. Also, let us write $dx_I = dx_{i_1} \wedge \ldots \wedge dx_{i_k}$, $1 \leq i_1 \leq \ldots \leq i_k \leq n$, where we say that $|I| = k$. The collection of wedge products like this forms a basis for $\Lambda^k(T_q^*M)$. Thus, if $\omega \in \Omega^k(M)$, then $\omega_x = \sum |I| = k a_I dx_I$, where $a_I \in C^\infty(M)$. $\square$

One major fact regarding forms is that we can differentiate them. The differentiation operator is called the **exterior derivative** and is characterized in the following theorem.

Theorem 6.3.
For every manifold M, there is a unique operator

$$d_M : \Omega(M) \to \Omega(M)$$

with the following properties :
(1) $d_M(\Omega^k(M)) \subset \Omega^{k+1}(M)$
(2) d_M is $\mathbb{R}$-linear
(3) $d_M f = df$ for all $f \in C^\infty(M)$
(4) d_M is local : For all open sets U and all $\omega \in \Omega(M)$, $(d_M\omega)\Big|_U = d_U(\omega\Big|_U)$
(5) $d_M(\omega \wedge \eta) = (d_M\omega) \wedge \eta + (-1)^k \omega \wedge (d_M\eta)$ for $\omega \in \Omega^k(M), \eta \in \Omega^l(M)$
(6) $d_M \circ d_M = 0$.

Proof. First we show that d_M is unique. Suppose that d_M exists. Fix a coordinate chart $(U,(x_1,\ldots,x_m))$. Then for all $\alpha \in \Omega^k(M)$, $\alpha\Big|_U = \sum_{|I|=k} a_I dx_I$, where $a_I \in C^\infty(U)$. Then

$$\begin{aligned} d\alpha\Big|_U &= d(\alpha\Big|_U) \\ &= d_U(\sum a_I dx_I) \\ &= \sum_I (d_U a_I \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_n} + \ldots) \\ &= \sum_I d_U a_I \wedge dx_I \end{aligned}$$

since $d_U^2 = 0$. Now, for all $f \in C^\infty(M)$,

$$df = \sum \frac{\partial f}{\partial x_j} dx_j.$$

So $da_I \wedge dx_I = \sum_j \frac{\partial a_I}{\partial x_j} dx_j \wedge dx_I$, which means that $d\alpha\Big|_U = \sum_{I,j} \frac{\partial a_I}{\partial x_j} dx_j \wedge dx_I$; i.e., $d\alpha\Big|_U$ is independent of coordinate chart, which says in turn that $d\alpha$ is uniquely defined. Thus, d is unique, provided that it exists.
But what about existence? For every coordinate patch $(U,(x_1,\ldots,x_n))$, define $d_U : \Omega^k(U) \to \Omega^{k+1}$ by $d_U(\sum a_I dx_I) = \sum da_I dx_I$. Note in particular that we define $d_U a = da$ if $a \in C^\infty(M)$ and that $d_U(dx_I) = 0$.
With this definition, let's check (1) $\to$ (6). First, (1),(2), and (3) follow almost directly by our definition of d, and are left to the reader.
(4) If $f \in C^\infty(U)$ and $V \subset U$ is open, then

$$d_U(f\Big|_V) = d(f\Big|_V) = df\Big|_V = (d_U f)\Big|_V.$$

(5) We check in coordinates: Let $I = (i_1,\ldots,i_k)$, $J = (j_i,\ldots,j_s)$, $i_r \neq j'_r$ $\forall r, r'$. Then

$$\begin{aligned} d_U(a_I dx_I \wedge b_J dx_J) &= d_U(a_I b_J dx_I \wedge dx_J) \\ &= \sum_j \frac{\partial}{\partial x_j}(a_I b_J) dx_j \wedge dx_I \wedge dx_J \\ &= \sum_j \frac{\partial a_I}{\partial x_j} b_J dx_j \wedge dx_I \wedge dx_J + \frac{\partial b_J}{\partial x_j} a_I dx_j \wedge dx_I \wedge dx_J \end{aligned}$$

$$\begin{aligned}
&= (\sum_j \frac{\partial a_I}{\partial x_j} dx_j \wedge dx_I) \wedge b_J dx_J + (-1)^k a_I dx_I \sum_j \frac{\partial b_J}{\partial x_j} b_J dx_j \wedge dx_J \\
&= d(a_I dx_I) \wedge (b_J) dx_J + (-1)^k (a_I dx_I) \wedge d(b_J dx_J).
\end{aligned}$$

(6) One again, in coordinates, we have

$$\begin{aligned}
d_U(d_U(a_I dx_I)) &= d_U(\sum_j \frac{\partial a_I}{\partial x_j} dx_j \wedge dx_I) \\
&= \sum_{i,j} \frac{\partial^2 a_I}{\partial x_i \partial x_j} dx_i \wedge dx_j \wedge dx_I \\
&= 0
\end{aligned}$$

since mixed partials commute and since $dx_i \wedge dx_j = -dx_j \wedge dx_i$. $\square$

Definition 6.4. Exterior Derivative
The operator described in the previous theorem is called the **exterior derivative**.

As a remark, note that $\omega : M \to \Lambda^k(T * M)$ is smooth if and only if for any smooth vector fields $X_1, \ldots, X_k \in \Gamma(TM)$, $\omega(X_1, \ldots, X_k)$ is a smooth function on M.

Now, since d is unique and local, we usually drop the subscript M and just write d. Also, keep in mind that for $f \in C^\infty(U)$, we have (locally)

$$df = \sum_i \frac{\partial f}{\partial x_i} dx_i.$$

Furthermore, note that if $\dim M = m$, then $\lambda^n(T_q^* M) = 0$ for all $q \in M$ if $n > m$; hence, $d\omega = 0$ for every $\omega \in \Omega^m(M)$.
At this point, the reader might ask where we are headed with this discussion. One answer to this question is that forms are important for us for two reasons: The first is that they will lead to the theory of integration, and in particular to Stokes's Theorem, which we will see in the next chapter. Also, however, the properties of d lead naturally to the de Rham cohomology groups of a manifold, which are a certain collection of groups associated with the forms on a manifold. The next few sections will develop the necessary machinery to talk about both of these topics. Specifically, we need:
1) Pull-backs: If $f : M \to N$ is smooth, it induces a linear map $f^* : \Omega(N) \to$

$\Omega(M)$.
2) Contractions: Given a vector field, we can define a linear map that takes k forms to $k-1$ forms.
3) Lie derivatives of forms: This is another way to differentiate forms which is related to contractions and the exterior derivative.
Armed with these concepts, we can then begin to discuss de Rham cohomology and in particular find the cohomology groups for $\mathbb{R}^n$ (the Poincare Lemma). We begin with pull-backs.

6.2 Pull-backs of Differential Forms

Let $f : M \to N$ be a smooth map. For all $q \in M$, we have $df_q : T_qM \to T_qN$, which in turn induces a map of dual spaces $(df_q)^* : T^*_{f(q)}N \to T^*_qM$. In turn, this map of dual spaces induces a map of exterior algebras $\Lambda^k((df_q)^*) : \Lambda^k(T^*_{f(q)}N) \to \Lambda^k(T^*_qM)$, for all k. In addition, we have, for all $\nu \in \Lambda^k(T^*_{f(q)}N)$,

$$\Lambda^k(df_q^*)(\nu)(v_1, \ldots, v_k) = \nu(df_q(v_1), \ldots, df_q(v_k)),$$

for all $v_1, \ldots, v_k \in T_qM$. Thus, we see that $f : M \to N$ defines a map $f^* : \Omega^k(N) \to \Omega^k(M)$ which is given by

$$(f^*\nu)(v_1, \ldots, v_k) = v_{f(q)}(df_q(v_1), \ldots, df_q(v_k)).$$

This induced map of forms f^* is called the **pull-back** . As we will see later on, it actually gives rise to our familiar change of coordinates formulae from calculus. For now, however, we'll prove a couple of lemmas which will tell us how to calculate pull-backs of forms, which we'll have to do from time to time throughout the next few sections. In particular, note that if $g \in C^\infty(M)$, then $f^*g = g \circ f$. In addition, the first lemma is automatic from the way we've defined the pull-back.

Lemma 6.5.
For all forms ω and μ in $\Omega(M)$, $f^*(\omega \wedge \mu) = f^*(\omega) \wedge f^*(\mu)$.

Lemma 6.6.
Let $f : M \to N$ be a map, and suppose that $\omega \in \Omega(M)$. Then

$$f^*d(\omega) = d(f^*\omega).$$

Proof. First consider the case that ω is a 0-form (i.e., $\omega \in C^\infty(M)$). Now, on $\Omega^0(N)$, $f^*\omega = \omega \circ f$. Then

$$\begin{aligned}
(f^*(d\omega))q(v) &= (d\omega)_{f(q)}((df)_q(v)) \\
&= (d\omega_{f(q)} \circ df_q)(v) \\
&= d(\omega \circ f)_q(v) \\
&= d(f^*\omega)_q(v).
\end{aligned}$$

Next, let $(U, (x_1, \ldots, x_n))$ be any coordinate chart on N and $\omega \in \Omega^k(N)$. Then $\omega\Big|_U = \sum_{|I|=k} a_I dx_I$. Thus, $d\omega\Big|_U = \sum_I da_I \wedge dx_I$, so that from the first case, we have

$$\begin{aligned}
f^*(d\omega\Big|_U) &= f^*(\sum da_I \wedge dx_I) \\
&= \sum f^*(da_I) \wedge f^*(dx_I) \\
&= \sum f^*(da_I) \wedge d(f^*x_{i_1}) \wedge \ldots \wedge d(f^*x_{i_n}).
\end{aligned}$$

Also, we have

$$\begin{aligned}
d(f^*\omega\Big|_U) &= d(f^*(\sum a_I dx_I)) \\
&= \sum f^*a_I f^*(dx_{i_1}) \wedge \ldots \wedge (dx_{i_n}) \\
&= \sum df^*(a_I) d(f^*x_{i_1}) \wedge \ldots \wedge d(f^*x_{i_n}).
\end{aligned}$$

Then again, by the first case, we see that

$$f^*(d\omega\Big|_U) = df^*(\omega\Big|_U).$$

□

Here's an example.

Example 6.7.
Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$(r, \theta) \mapsto (r\cos\theta, r\sin\theta).$$

Let's calculate the pull-back of the volume form $dx \wedge dy$ on $\mathbb{R}^2$ by this map.

$$\begin{aligned}
f^*(dx \wedge dy) &= d(f^*x) \wedge d(f^*y) \\
&= (\cos\theta dr - r\sin\theta d\theta) \wedge (\sin\theta dr + r\cos\theta d\theta) \\
&= (\cos\theta \cdot r\cos\theta dr \wedge d\theta + (-r\sin\theta)\sin\theta d\theta \wedge dr) \\
&= r dr \wedge d\theta.
\end{aligned}$$

□

6.3 Contractions

Now we'll talk some about contractions. Let V be a finite-dimensional real vector space with $u \in V$. Let $\eta \in \Lambda^k(V^*)$. Define the **contraction of** u **with** η to be the following element of $\Lambda^{k-1}(V^*)$:

$$(\iota(u)\eta)(v_1, \ldots, v_{k-1}) = \eta(u, v_1, \ldots, v_{k-1}).$$

Thus, given $u \in V$, $\iota(u)$ defines a map from $\Lambda^k(V^*) \to \Lambda^{k-1}(V^*)$. Note a special case of this definition:
If $\eta \in \Lambda^1(V^*) \simeq V^*$, then $\iota(u)\eta = \eta(u)$. Thus, $\iota(u) : V^* \to \mathbb{R}$.

Example 6.8.
Suppose $l_1, l_2 \in V^*$, so that $l_1 \wedge l_2 \in \Lambda^2(V^*)$. Then

$$\begin{aligned}
(\iota(u)(l_1 \wedge l_2))(v) &= (l_1 \wedge l_2)(u, v) \\
&= (l_1 \wedge l_2)(u \wedge v) \\
&= l_1(u)l_2(v) - l_1(v)l_2(u)
\end{aligned}$$

$$= l_1(u)l_2(v) - l_1(v)l_2(u)$$

That is, $\iota(u)(l_1 \wedge l_2) = l_1(u)l_2 - l_2(u)l_1$. □

Now for the following lemma, which will tell us exactly how to calculate contractions.

Lemma 6.9.
If $l_1, \ldots, l_k \in V^*$, $u \in V$, then

$$\iota(u)(l_1 \wedge \ldots \wedge l_k) = \sum_{j=1}^{k} (-1)^{j-1} (\iota(u)l_j)(l_1 \wedge \ldots \wedge \hat{l_j} \wedge \ldots \wedge l_k).$$

Proof.

$$\begin{aligned}
(\iota(u)l_1 \wedge \ldots \wedge l_k)(v_1, \ldots, v_{k-1}) &= \det \begin{pmatrix} l_1(u) & l_1(v_1) & \ldots & l_1(v_{k-1}) \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ l_k(u) & l_k(v_1) & \ldots & l_k(v_{k-1}) \end{pmatrix} \\
&= \sum_{j=1}^{k} (-1)^{j-1} l_j(u) \det \\
&= \sum_{j=1}^{k} (-1)^{j-1} l_j(u) (l_1 \wedge \ldots \wedge \hat{l_j} \wedge \ldots \wedge l_k(v_1, \ldots, v_{k-1}))
\end{aligned}$$

□

Corollary 6.9.1.
If $\alpha \in \Lambda^r(V^*)$, $\beta \in \Lambda^s(V^*)$, $u \in V$, *then*

$$\iota(\alpha \wedge \beta) = (\iota(u)\alpha) \wedge \beta + (-1)^r \alpha \wedge (\iota(u)\beta).$$

Proof. It's enough to consider $\alpha = l_1 \wedge \ldots \wedge l_r$, $\beta = l_{r+1} \wedge \ldots \wedge l_{r+s}$. This we leave to the reader. □

Now let's do another example, this time using forms and vector fields. The point of contracting forms and vector fields is, analogous to what we defined above, $\iota(X) : \Omega^k(M) \to \Omega^{k-1}(M)$. To calculate a contraction in this case, we simply make use of the previous lemma. Let's see an explicit case of how this works.

Example 6.10.
Let W be the vector field on $\mathbb{R}^3$ given by $x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}$, and let $\omega = dx \wedge dy \wedge dz$, the volume form on $\mathbb{R}^3$. Then

$$\begin{aligned} \iota(W)\omega &= \iota(\omega)dx \wedge dy \wedge dz \\ &= dx(\omega)dy \wedge dz - dy(\omega)dx \wedge dz + dz(\omega)dx \wedge dy \\ &= xdx \wedge dy - ydy \wedge dz + zdx \wedge dy \end{aligned}$$

□

Thus, the only thing we need to know to compute contractions (locally, at least) is to remember how functionals dx_i act on vector fields and to follow the "alternating rule" for computing contractions outlined in the previous lemma.

6.4 Lie Derivatives of Forms

Finally, we come to Lie derivatives, the final piece of machinery that we'll need to begin discussing de Rham cohomology.

Definition 6.11.
Let X be a vector field on M, $\omega \in \Omega^k(M)$. Let ϕ_t denote the local flow of X. The **Lie derivative of ω with respect to X** is

$$(L_X\omega)_q = \frac{d}{dt}\bigg|_t (\phi_t^*\omega)_q.$$

As with Lie derivatives of vector fields, there is an explicit formula for calculating Lie derivatives of forms, known as **Cartan's formula**. Before we state and prove this formula, we need to establish a couple of lemmas. Before we move on to state these lemmas, however, let's figure out what $L_X f$ should be, where $f \in C^\infty(M)$.

$$\begin{aligned}
(L_X f)_q &= \frac{d}{dt}\Big|_{t=0} (\phi_t^* f)_q \\
&= \frac{d}{dt}\Big|_{t=0} (f \circ \phi_t)(q) \\
&= X(f)(q) = \\
&= df_q(X(q)) \\
&= (\iota(X)\, df)_q
\end{aligned}$$

Now for the first lemma.

Lemma 6.12.
Let X be a vector field on M. L_X is a derivation on $\Omega(M)$ which commutes with d.

Proof. First, we need to unravel the statement of the theorem. There are three things we need to prove:
1) L_X is $\mathbb{R}$-linear.
2) $L_X(\omega \wedge \eta) = (L_X \omega) \wedge \eta + \omega \wedge (L_X \eta)$
3) $L_X(d\omega) = d(L_X \omega)$.
(1) is immediate since pull-backs and differentiation are linear, so let's look at (2).

$$\frac{d}{dt}\Big|_{t=0} (\phi_t^*(\omega \wedge \eta) = \frac{d}{dt}\Big|_{t=0} (\phi_t^* \omega) \wedge (\phi_t^* \eta).$$

As $\wedge$ is bilinear,

$$\begin{aligned}
\frac{d}{dt}\Big|_{t=0} (\phi_t^* \omega) \wedge (\phi_t^* \eta) &= \frac{d}{dt}\Big|_{t=0} (\phi_t^* \omega) \wedge (\phi_0^* \eta) + (\phi_0^* \omega) \frac{d}{dt}\Big|_{t=0} (\phi_t^* \eta) \\
&= (L_X \omega) \wedge \eta + \omega \wedge (L_X \eta).
\end{aligned}$$

One can also check that this statement is true in coordinates.
Now for (3).

$$\begin{aligned}
L_X(d\omega) &= \frac{d}{dt}\Big|_{t=0} (\phi_t^*(d\omega)) \\
&= \frac{d}{dt}\Big|_{t=0} d(\phi^* \omega)
\end{aligned}$$

$$
\begin{aligned}
&= d\left(\frac{d}{dt}\Big|_{t=0}\phi_t^*\omega\right) \\
&= d(L_X\omega).
\end{aligned}
$$

□

Lemma 6.13. *Let X be a vector field on M, $Q = d \circ \iota(X) + \iota(X) \circ d$. Then Q is a derivation on $\Omega^k(M)$ that commutes with d.*

Proof.

$$
\begin{aligned}
Q \circ d &= d\, \iota(X)\, d + \iota(X)\, d\, d \\
&= d\, \iota(X)\, d \\
&= d\, d\, \iota(X) + d\, \iota(X)\, d \\
&= d \circ Q
\end{aligned}
$$

Thus, Q and d commute. Now we need to check that Q is a derivation. Accordingly, let $\omega \in \Omega^k(M)$, $\eta \in \Omega^l(M)$. Then

$$
\begin{aligned}
Q(\omega \wedge \eta) &= d(\iota(X)(\omega \wedge \eta)) + \iota(X)(d(\omega \wedge \eta)) \\
&= d((\iota(X)\omega) \wedge \eta + (-1)^k \omega \wedge (\iota(X)\eta)) + \iota(X)((d\omega) \wedge \eta + (-1)^k \omega \wedge (d\eta)). \\
&= d(\iota(X)\omega) \wedge \eta + \omega \wedge d(\iota(X)\eta) + (\iota(X)d\omega) \wedge \eta + \omega \wedge (\iota(X)d\eta) \\
&= Q(\omega) \wedge \eta + \omega \wedge Q(\eta)
\end{aligned}
$$

□

The next theorem gives us an explicit way to calculate Lie derivatives of forms utilizing contractions and exterior differentiation.

Theorem 6.14. Cartan's Formula
Suppose that X is a vector field on M and $\omega \in \Omega(M)$. Then we have **Cartan's Formula** *:*

$$L_X\omega = d(\iota(X)\omega) + \iota(X)d\omega.$$

Proof. It's enough to prove

$$(L_X\omega)\Big|_U = (d\iota(X)\omega + \iota(X)d\omega)\Big|_U$$

for any coordinate chart U. Thus, we may further assume that $\omega = a_I dx_I = a_I dx_{i_1} \wedge \ldots \wedge dx_{i_n}$. Both L_X and $Q = d\iota(X) + \iota(X)d$ are derivations that commute with d, so it's enough to prove the above equation for functions, which we did. □

Exercise 6.1.
Give examples of non-zero 2-forms ω and μ on $\mathbb{R}^4$ such that $\omega \wedge \omega = 0$ and $\mu \wedge \mu \neq 0$.

Exercise 6.2.
Consider polar coordinates (r, θ) on $\mathbb{R}^2$. The "function" θ is defined up to a constant. Show that $d\theta$ is a well-defined 1-form on $\mathbb{R}^2 - \{0\}$ and that

$$d\theta = \frac{1}{x^2 + y^2}(x\ dy - y\ dx).$$

Exercise 6.3.
(1) Let $\omega = x_1 dx_2 \wedge dx_3 + x_2 dx_3 \wedge dx_1 + x_3 dx_1 \wedge dx_2$. Compute $d\omega$.
(2) Compute

$$\iota(\sum_{i=1}^{3} x_i \frac{\partial}{\partial x_i}) dx_1 \wedge dx_2 \wedge dx_3.$$

(3) Compute $L_X(dx_1 \wedge dx_2 \wedge dx_3)$, where $X = \sum_{i=1}^{3} x_i \frac{\partial}{\partial x_i}$.

Exercise 6.4.
Consider $k : \mathbb{R}^2 \to \mathbb{R}^2$ given by $(u, v) \mapsto (u^2 + 1, uv)$. Compute $k^*((xy - y)dx \wedge dy)$.

Exercise 6.5.
Let X and Y be vector fields and α a 1-form on a manifold M. Prove that
1) $L_X(\iota(Y)\alpha) = (L_X\alpha) + \alpha(L_X Y)$.
2) Using (1), show that $d\alpha(X, Y) = X(\alpha(Y)) - Y(\alpha(X)) - \alpha([X, Y])$.

Exercise 6.6.
On $\mathbb{R}^3$, consider the 2-form

$$\alpha = y dx \wedge dz + \sin{(xy)} dx \wedge dy + e^x dy \wedge dz,$$

and let

$$X = z\frac{\partial}{\partial y}.$$

Compute $d\alpha$, $\iota(X)\alpha$ and $L_X\alpha$.

Exercise 6.7.
For any vector field X and any one form ω, define a 1-form by the Leibniz type-rule by

$$X(\omega(Y)) = (L_X\omega)(Y) + \omega([X, Y]).$$

From this definition, deduce Cartan's formula.

6.5 de Rham Cohomology

Introduction

Now we're ready to talk about de Rham cohomology. First, we introduce the following definitions.

Definition 6.15. Closed and Exact Forms
$\omega \in \Omega^k(M)$ is **closed** if it satisfies the differential equation $d\omega = 0$.
$\omega \in \Omega^k(M)$ is **exact** it satisfies the differential equation if $\omega = d\eta$ for some $\eta \in \Omega^{k-1}(M)$.

Note that since $d^2 = 0$, **every exact form is closed**, but the converse need not be true. To some extent, de Rham cohomology is a measure of the extent to which every closed form is exact.
Since $d^2 = 0$, we have that

$$\mathrm{Im}\{d : \Omega^k(M) \to \Omega^{k+1}(M)\} \subset \ker\{d : \Omega^{k+1}(M) \to \Omega^{k+2}(M)\}.$$

Thus, it makes sense to consider

$$H^k(M) = \ker\{d : \Omega^{k+1}(M) \to \Omega^{k+2}(M)\}/\mathrm{Im}\{d : \Omega^k(M) \to \Omega^{k+1}(M)\};$$

that is, it makes sense to consider the collection of closed k-forms modulo the collection of exact k-forms on a manifold. These are the **de Rham Cohomology groups**. Note that, in fact, these are groups since closed forms and exact forms both form vector spaces over $\mathbb{R}$.

Definition 6.16. de Rham Cohomology
Define $H^k(M)$ to be the closed k-forms modulo the exact k-forms. The groups $H^k(M)$ are called the **de Rham Cohomology groups of M.**

Now, let's make a couple of observations:
First, $H^0 = \{f \in \Omega^0(M) : df = 0\}$ is the collection of locally constant functions. Thus, if M is connected, we have immediately that $H^k(M) \simeq \mathbb{R}$. Generally, we have that $H^0(M)$ is a direct product of copies of $\mathbb{R}$, one for each connected component of M. Next, we can see immediately that $H^n(M) = \{0\}$ if $n > \dim M$.

de Rham Cohomology as a Ring

Proposition 6.17.
The space of closed forms forms a ring, where the multiplication is given by the wedge product. The space of exact forms is an ideal in that ring.

Proof. Since wedge product is associated, we need only show that the wedge product of two closed forms is closed. Thus, let ω and μ be two closed forms; then

$$d(\omega \wedge \mu) = d\omega \wedge \mu + (-1)^k \omega \wedge d\mu = 0.$$

Now, to show that the space of exact forms is a subring, we'll prove that the wedge of two exact forms is exact; the other necessary properties are immediate. Thus, let $d\eta = \mu$ and $d\omega = \nu$ Then

$$\nu \wedge \mu = d(\omega \wedge \mu).$$

Finally, suppose that ω is a closed form and that $\mu = d\eta$. Then $\omega \wedge \mu$ is exact:

$$\omega \wedge \mu = d(\omega \wedge \eta).$$

Similarly, we see that $\mu \wedge \omega$ is exact. Therefore, the exact forms are a two-sided ideal in the ring of closed forms with the multiplicative structure induced by the wedge product. □

Corollary 6.17.1.
de Rham cohomology is a ring under wedge product.

Actually, we can say even more: de Rham cohomology is a graded $\mathbb{R}$-algebra.

de Rham Cohomology as a Functor

Now suppose that $f : M \to N$ is smooth, $\omega \in \Omega^k(N)$ is closed and $\mu \in \Omega^k(M)$ is exact. Then the fact that f^* commutes with d means that $f^*(\omega)$ is closed on M and $f^*(\mu)$ is exact on M. That is, f^* induces an $\mathbb{R}$-linear map

$$f^* : H^k(N) \to H^k(M).$$

This hints at the following proposition:

Proposition 6.18.
The k^{th} de Rham cohomology is a contravariant functor from the category of differentiable manifolds to the category of real vector spaces.

Proof. Let T be the map that takes M to $H^k(M)$ and a smooth map $f : M \to N$ to its pull-back f^*. Then we know that f^* is a linear map from $H^k(N)$ to $H^k(M)$. Furthermore, it is easy to see that $T(1_M) = 1_{H^k(M)}$. Finally, if $f : M \to N$ and $g : N \to Q$ are smooth maps, then $T(fg) = (fg)^* : H^k(Q) \to H^k(M)$; since $(fg)^* = g^* f^*$, however, it follows that $T(fg) = g^* f^*$. □

Corollary 6.18.1.
If $f : M \to N$ is a diffeomorphism, then $f^* : H^k(N) \to H^k(M)$ is an isomorphism.

Calculating Cohomology Groups for $\mathbb{R}^n$

Now we'll calculate de Rham cohomology groups for $\mathbb{R}^n$. The next theorem asserts that the cohomology groups in this case are very simple. Here's where we'll use Lie derivatives and contractions.

Theorem 6.19. Poincare Lemma
Let $B_R(0) \subset \mathbb{R}^n$ be a ball of radius R centered at the origin. Then $H^k(B_R(0)) = \{0\}$ for $k > 0$. In particular, this says that every closed form is exact in $\mathbb{R}^n$.

Proof. Let $M = B_R(0)$. We will construct $h_k : \Omega^k(M) \to \Omega^{k-1}$, $k \geq 1$, such that

$$h_{k+1} \circ d + d \circ h_k = id_{\Omega^k(M)}.$$

Then if $\omega \in \Omega^k(M)$ and $d\omega = 0$, we'll have

$$\omega = (h_{k+1} \circ d + d \circ h_k)\omega = d(h_k \circ \omega);$$

i.e., every closed form is exact. So if we construct these maps, we're done. Visually, we want

$$\begin{tikzcd}
\Omega^0(M) \arrow[r, "d"] \arrow[d, equal] & \Omega^1(M) \arrow[r, "d"] \arrow[d, equal] \arrow[dl, "h_1"'] & \Omega^2(M) \arrow[r, "d"] \arrow[d, equal] \arrow[dl, "h_2"'] & \cdots \\
\Omega^0(M) \arrow[r, "d"] & \Omega^1(M) \arrow[r, "d"] & \Omega^2(M) \arrow[r, "d"] & \cdots
\end{tikzcd}$$

Thus, we turn to the construction of h_k. First, define a linear map $\alpha_k : \Omega^k(M) \to \Omega^k(M)$ by

$$\alpha_k(a dx_I) = (\int_0^1 a(tx)t^{k-1}dt)dx_I.$$

Let $X = \sum x_i \frac{\partial}{\partial x_i}$. We now have the following claims.
Claim I: $\alpha_k \circ L_X = id_{\Omega^k}(M)$.
Claim II: $\alpha_k \circ d = d \circ \alpha_{k-1}$.
For a moment, let us assume that these two claims are true. Then

$$\begin{aligned}
id_{\Omega^k(M)} &= \alpha_k \circ L_X \\
&= \alpha_k \circ (\iota(X) \circ d + d \circ \iota(X)) \\
&= (\alpha_k \circ \iota(X)) \circ d + d \circ (\alpha_{k-1} \circ \iota(X)) \\
&= h_{k+1} + h_k.
\end{aligned}$$

Therefore, if we establish the two claims, we're done.
Claim I:
Fix $a \in C^\infty(M)$. Then

$$L_X(a) = (\sum_i x_i \frac{\partial}{\partial x_i})(a) = \sum_i x_i \frac{\partial a}{\partial x_i}.$$

In addition,

$$L_X(dx_j) = d(L_X x_j) = d(\sum_i (x_i \frac{\partial}{\partial x_i})(x_j)) = dx_j.$$

Thus,

$$\begin{aligned} L_X(adx_I) &= L_X(adx_{i_1} \wedge \ldots \wedge dx_{i_k} \\ &= L_X(a)dx_I + a(L_X dx_{i_1}) \wedge \ldots \wedge dx_{i_n} + \ldots + adx_{i_1} \wedge \ldots \wedge (L_X dx_{i_n}) \\ &= (\sum_i x_i \frac{\partial a}{\partial x_i} + ka)dx_I. \end{aligned}$$

Therefore,

$$\alpha_k(L_X(adx_I)) = (\int_0^1 t^{k-1}(\sum_i tx_i \frac{\partial a}{\partial x_i}(tx) + ka(tx))dt)dx_I.$$

But the integrand here is just

$$t^k(\sum_i tx_i \frac{\partial a}{\partial x_i} + ka)(tx) = \frac{d}{dt}(t^k a(tx)).$$

Thus,

$$\int_0^1 \frac{d}{dt}(t^k a(tx))dt = t^k \cdot a(tx)\Big|_0^1 = ax - 0.$$

Thus, $\alpha_k \circ L_X = id_{\Omega^k(M)}$.
Claim II: Note that

$$\begin{aligned} \alpha_k(d(a \cdot dx_I)) &= \alpha_k(\sum_i \frac{\partial a}{\partial x_i} dx_i \wedge dx_I \\ &= \sum_i (\int_0^1 t^k \frac{\partial a}{\partial x_i}(tx)dt)dx_i \wedge dx_I \end{aligned}$$

Thus,

$$\begin{aligned} d(\alpha_{k-1}(adx_I)) &= d(\int_0^1 t^{k-2}a(tx)dt)dx_I \\ &= \sum_i \frac{\partial}{\partial x_i}(\int_0^1 t^{k-2}a(tx)dt)dx_i \wedge dx_I \\ &= \sum_i (\int_0^1 t^{k-2}t\frac{\partial}{\partial a}x_i(tx)dt)dx_I \end{aligned}$$

$\square$

We will also calculate the cohomology groups for S^1 after we have defined integration of forms.

Exercise 6.8.
In $\mathbb{R}^3$, the standard inner product defines an isomorphism $(\mathbb{R}^3)^* \simeq \mathbb{R}^3$, which in turn induces an isomorphism of vector spaces $A : \Gamma(T\mathbb{R}^3) \to \Omega^1(\mathbb{R}^3)$. The standard volume form $\mu = dx_1 \wedge dx_2 \wedge dx_3$ defines an isomorphism $\mathbb{R}^3 \mapsto \Lambda^2((\mathbb{R}^3)^*)$ by $v \mapsto \iota(v)\mu$, which also induces an isomorphism $B : \Gamma(T\mathbb{R}^3) \mapsto \Omega^2(\mathbb{R}^3)$ given by $B(X) = \iota(X)\mu$. Finally, the map $C : C^\infty(\mathbb{R}^3) \to \Omega^3(\mathbb{R}^3)$ given by $C(f) = f\mu$ is also an isomorphism. (Check these facts!)
(a) Show that the standard vector calculus notions of div, grad, and curl can be defined as
• $\text{grad}(f) = A^{-1}(df)$.
• $\text{curl}(f) = B^{-1}(d(A(f)))$.
• $\text{div}(f) = C^{-1}(d(B(f)))$.
(b) Prove that in $\mathbb{R}^3$,
• $\text{curl}(f) = 0$ if and only if $f = \text{grad}(f)$ for some $f \in C^\infty(\mathbb{R}^3)$
• $\text{div}(f) = 0$ if and only if $f = \text{curl}(g)$ for some $g \in \Gamma(T\mathbb{R}^3)$

Exercise 6.9.
Let $\mathbb{R}^3$ be standard global coordinates on $\mathbb{R}^3$. Consider the 2-form

$$\omega_c = y \cdot dx \wedge dz + (-c)x \cdot dy \wedge dz,$$

where $c \in \mathbb{R}$. Determine all values of c for which there is a form η such that $d\eta = \omega_c$.

7 Integration of Differential Forms

Now we discuss integration of forms on manifolds. This is a natural generalization of our vector calculus notions of integration, and we'll especially be interested in Stokes's Theorem, which will generalize Green's Theorem in the plane and the divergence theorem (a.k.a. Gauss's Theorem). Our first task, however, will be to define integration of forms on arbitrary manifolds and to prove that, under certain conditions, integration is well-defined linear map from the collection of forms to the real line.

7.1 Integration of Differential Forms

Definition 7.1.
Let $\Omega^n_C(M)$ denote the collection of compactly supported n-forms on M. Similarly, we denote the collection of compactly supported smooth functions by $C^\infty_C(M)$.

If $\omega \in \Omega^n_C(\mathbb{R}^n)$, then $\omega = a(x)dx_1 \wedge \cdots \wedge dx_n$ for some $a \in C^\infty_C(\mathbb{R}^n)$. We **define**

$$\int_{\mathbb{R}^n} \omega = \int_{\mathbb{R}^n} a(x).$$

Note that the definition does indeed make sense since $a(x)$ is compactly supported (i.e., the integral is finite).
Next, recall from advanced calculus that if $U \subset \mathbb{R}^n$ is open, $f \in C^\infty_C(M)$, and $f : V \to U$ is a diffeomorphism, then

$$\int_V (f \circ F)(y) \cdot |\det dF(y)| dy = \int_{U=F[V]} f(x)dx \tag{1}$$

This is merely the generalization of u-substitution (i.e., change of variables) to multivariable calculus.
Now, suppose additionally that $\det dF(y) > 0$ for all y. Then (1) can be expressed as follows:

$$\begin{aligned}\int_V (f \circ F) \cdot \det |dF(y)| dy &= \int_V (F^* f)(\det dF(y)) dy_1 \wedge \ldots \wedge dy_n \\ &= \int_V F^*(f dx_1 \wedge \ldots \wedge dx_n).\end{aligned}$$

This gives the **change of variables formula for forms**:

$$\int_V F^*\omega = \int_{F[V]} \omega,$$

where $F : V \to F[V]$ is a diffeomorphism, ω is compactly supported, and $\det dF(y) > 0$ for all $y \in V$.

For this formula to make sense, however, we must be able to ensure that $\det dF(y) > 0$. The problem is that, while we may be able to do this within a single coordinate chart, the support of ω might lie across several coordinate charts, and there is no guarantee that the sign of the determinant will be preserved from one coordinate chart to another. Thus, we make the following definition.

Definition 7.2. Orientability
A manifold M is **orientable** if there is an atlas $\{(U_\alpha, \phi_\alpha)\}$ such that for all α and β with $U_\alpha \cap U_\beta \neq \emptyset$,

$$\det(d(\phi_\beta \circ \phi_\alpha^{-1})) > 0.$$

The choice of such an atlas is an **orientation**, and an **oriented manifold** is a manifold together with an orientation.

Example 7.3.
$\mathbb{R}^n$ with the identity map as a chart is an oriented manifold. $\square$

Example 7.4.
The Mobius band is non-orientable. A proof is outlined in Rudin's undergraduate analysis book. $\square$

As it turns out, orientable is the condition we need to make the following important theorem go through:

Theorem 7.5.
Let M be an oriented n-dimensional manifold. There exists a unique linear map (integration) $\int_M : \Omega^n_C(M) \to \mathbb{R}$, given by

$$\omega \mapsto \int_M \omega,$$

such that if (U, ϕ) is a coordinate chart and $\omega \in \Omega^n_C(U)$, then

$$\int_M \omega = \int_{\phi(U)} (\phi^{-1})^* \omega = \int_{\mathbb{R}^n} (\phi^{-1})^* \omega.$$

Proof. We need to check that integration of forms is well-defined and unique, since linearity follows from familiar properties of the Riemann (or Lebesgue) integrals from calculus. We do this in two steps.

Step I: First, we check that if the support of ω is in U and (U, ϕ), (U, ψ) are two different charts defining the same orientation, then

$$\int_{\mathbb{R}^n} (\phi^{-1})^* \omega = \int_{\mathbb{R}^n} (\psi^{-1})^* \omega.$$

Hence $\int_M \omega$ will be well-defined (locally) and will be automatically unique. Now, $\phi^{-1} = \psi^{-1} \circ (\psi \circ \phi^{-1})$, so

$$\begin{aligned} (\psi^{-1})^* \omega &= (\psi^{-1} \circ (\psi \circ \phi^{-1}))^* \omega \\ &= (\psi \circ \phi^{-1})^* ((\psi^{-1})^* \omega) \end{aligned}$$

Then by (1), we have

$$\begin{aligned} \int_{\phi[U]} (\phi^{-1})^* \omega &= \int_{\phi[U]} (\psi \circ \phi^{-1})^* ((\psi^{-1})^* \omega) \\ &= \int_{\psi[U]} (\psi^{-1})^* \omega. \end{aligned}$$

Step II: Now we do the general case. Here, the proof relies on tha fact that partitions of unity exist. Thus, let ω be an arbitrary compactly supported form on M. Let $\{(U_\alpha, \phi_\alpha)\}$ be an atlas on M giving it its orientation. Furthermore, let $\{p_\alpha\}$ be a parition of unity subordinate to $\{U_\alpha\}$. Note that since supp(ω) is compact, $p_\alpha \neq 0$ for only finitely many α. Now **define**:

$$\begin{aligned} \int_M \omega &= \sum_\alpha \int_{\phi_\alpha[U_\alpha]} (\phi_\alpha^{-1})^* (p_\alpha \omega) \\ &= \sum_\alpha \int_M p_\alpha \omega. \end{aligned}$$

Now, note that since $p_\alpha \neq 0$ for only finitely many α, we do not need to justify interchanging the summation and the integral. Thus, we need to show only that $\int_M \omega$ does not depend on the choice of atlas or on the choice of the partition of unity. Accordingly, suppose that $\{(V_\beta, \psi_\beta)\}$ is another atlas giving M its orientation and that $\{\tau_\beta\}$ is a partition of unity subordinate to $\{V_\beta\}$. By step 1,

$$\int_{\psi_\beta[U_\alpha \cap V_\beta]} (\psi_\beta^{-1})^* (\tau_\beta p_\alpha \omega) = \int_{\phi_\alpha[U_\alpha \cap V_\beta]} (\phi_\alpha^{-1})^* (\tau_\beta p_\alpha \omega).$$

Therefore,

$$\begin{aligned}
\sum_{\alpha} \int_{\phi_\alpha[U_\alpha]} (\phi_\alpha^{-1})^*(p_\alpha \omega) &= \sum_{\alpha,\beta} \int_{\phi_\alpha[U_\alpha]} (\phi_\alpha^{-1})^*(p_\alpha \tau_\beta \omega) \\
&= \sum_{\alpha,\beta} \int_{\phi_\alpha[U_\alpha \cap V_\beta]} (\phi_\alpha^{-1})^*(p_\alpha \tau_\beta \omega) \\
&= \sum_{\alpha,\beta} \int_{\psi_\beta[U_\alpha \cap V_\beta]} (\psi_\beta^{-1})^*(p_\alpha \tau_\beta \omega) \\
&= \sum_{\beta} \int_{\psi_\beta[V_\beta]} (\psi_\beta^{-1})^*(\tau_\beta \omega)
\end{aligned}$$

Thus, we see that

$$\int_M \omega$$

is well-defined. □

Example 7.6.
Suppose that $M = S^1$. Let θ be a coordinate chart on M; note that this is a homeomorphism from $S^1 - \{(1,0)\}$ to $(0, 2\pi)$. Then

$$\int_{S^1} \sin\theta \; d\theta = \int_0^{2\pi} \sin\theta \; d\theta.$$

□

If M is a manifold of dimension m, then a **top form on** M (or **volume form**) is an m-form. Here is an equivalent condition for being orientable.

Proposition 7.7.
An n-dimensional manifold is orientable if and only if there is a non-vanishing top form on M.

Proof. Let $\{(U_\alpha, \phi_\alpha)\}$ be an atlas giving M an orientation. Assume further that all U_α are connected, and let $\{p_\alpha\}$ be a partition of unity subordinate to $\{U_\alpha\}$. Define

$$\nu = \sum_\alpha p_\alpha(\phi_\alpha^*(dx_1 \wedge \ldots \wedge dx_n)).$$

We claim that $\nu_x \neq 0$ for all $x \in M$. To see this, fix $x \in M$. Then $p_\alpha(x) \neq 0$ for finitely many α, say $\alpha_1, \ldots \alpha_k$. Then

$$\begin{aligned}
(\phi_{\alpha_1}^{-1})^*\nu &= \sum_{i=1}^{k}(\phi_{\alpha_1}^{-1})^* p_{\alpha_i}\phi_{\alpha_i}^*(dx_1 \wedge \ldots \wedge dx_n) \\
&= \sum_{i=1}^{k}(p_{\alpha_i} \circ \phi_{\alpha_1}^{-1})[(\phi_{\alpha_i} \circ \phi_{\alpha_1}^{-1})^*(dx_1 \wedge \ldots \wedge dx_n)] \\
&= \sum_{i=1}^{k}(p_{\alpha_i} \circ \phi_{\alpha_1}^{-1})[\det d(\phi_{\alpha_i} \circ \phi_{\alpha_1}^{-1})]dx_1 \wedge \ldots \wedge dx_n \\
&\neq 0
\end{aligned}$$

since the determinants are always positive at x.
Now assume that ν is a non-vanishing top form. We want to produce an atlas with "positive" transition maps. So let $\{U_\alpha, \phi_\alpha\}$ be an atlas on M. Then $(\phi_\alpha^{-1})^*\nu = f_\alpha(x)dx_1 \wedge \ldots \wedge dx_n$, for some f_α, and the fact that ν is non-vanishing means that $f_\alpha \neq 0$. Thus, since U_α is connected, either $f_\alpha(x) > 0\ \forall x \in U_\alpha$ or $f_\alpha(x) < 0\ \forall x \in U_\alpha$. If $f_\alpha(x) > 0$, keep ϕ_α; otherwise, replace it with $T \circ \phi_\alpha$, where $T : \mathbb{R}^n \to \mathbb{R}^n$ is given by $(x_1, \ldots, x_n) \mapsto (-x_1, x_2, \ldots, x_n)$. $\square$

Exercise 7.1.
Suppose that M and N are orientable manifolds. Must it be true that $M \times N$ is orientable? Give a proof or a counterexample.

Exercise 7.2.
Show that TM is always orientable, regardless of whether or not M is.

Exercise 7.3.
(a) If $\mu \in \Lambda^n((\mathbb{R}^n)^*)$, prove that the map $\psi : \mathbb{R}^n \mapsto \Lambda^{n-1}(\mathbb{R}^n)$ is an isomorphism.
(b) Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is smooth and that a is a regular value of f. Show that $f^1(a)$ is an orientable manifold and that $\iota(\nabla f)$ gives $f^{-1}(a)$ an orientation. Here, of course, $\nabla f = \sum_i \frac{\partial f}{\partial x_i}\frac{\partial}{\partial x_i}$.

7.2 Stokes's Theorem

Definition 7.8. Regular Domain
Let M be a manifold of dimension n (without boundary). A subset $D \subset M$ is

a **regular domain** (or alternatively, a **domain with smooth boundary**) if for all $p \in D$, there is a coordinate chart $(U, \phi = (x_1, \ldots, x_n))$ such that

$$\phi(U \cap D) = \phi(U) \cap \{x \in \mathbb{R}^n : x_1 \leq 0\}.$$

Here, ϕ is said to be a chart **adapted to D**.

Example 7.9.
$M = \mathbb{R}^2$, D is the unit disk. Then $\phi(x_1, x_2) = (x_1 - \sqrt{1 - x_2^2}, x_2)$ works. □

Lemma 7.10.
Let D be a regular domain in M. Then ∂D is a submanifold of M that has codimension 1.

Proof. Suppose that (ϕ, U) and (ψ, V) are two charts adapted to D. Then $\psi \circ \phi^{-1} : \phi[U \cap V] \to \psi[U \cap V]$ sends $\phi[U \cap V] \cap \{x_1 \leq 0\}$ to $\psi[U \cap V] \cap \{x_1 \leq 0\}$. This restricts to a diffeomorphism $\phi[U \cap V] \cap \{x_1 = 0\} \to \psi[U \cap V] \cap \{x_1 = 0\}$ since in particular, it must map topological boundary to topological boundary. □

A couple of remarks:
First, note that

$$d(\psi \circ \phi^{-1})(0, x_2, \ldots, x_n) = \begin{pmatrix} a & 0 & \cdots & 0 \\ \vdots & & & \\ \vdots & & d(\psi \circ \phi^{-1})\Big|_{\{0\} \times \mathbb{R}^{n-1}} & \\ \vdots & & & \end{pmatrix}$$

where $a(x_2, \ldots, x_{n-1}) \in C^\infty(\mathbb{R}^{n-1})$ and $a > 0$.
Next, note that if (ϕ, U) and (ψ, V) are two charts adapted to D and $\det d(\phi \circ \psi^{-1}) > 0$, then $\det d(\phi \circ \psi^{-1})\Big|_{\{0\} \times \mathbb{R}^{n-1}} > 0$. Thus, if M is orientable, so is D.
In light of the previous fact, a natural question to ask is: Given an orientation of M, how do we orient ∂D? The following proposition will take care of this problem.

Proposition 7.11.
Let D be a regular domain of M and μ a non-vanishing top form. Then there is a vector field N defined on M near ∂D which points out of D. Moreover,

$$\nu = (\iota(N)\mu)\Big|_{\partial D}$$

is an orientation on ∂D.

Proof. We prove this fact in two steps.
Claim I: There is a vector field N defined on M near ∂D that points out of D.
If $D = \{x \in \mathbb{R}^n : x_1 \leq 0\}$, take $N = \frac{\partial}{\partial x_1}$. In general, cover D by adapted charts$\{(U_i, \phi_i\}$. On each U_i, there is $N_i \in \Gamma(TU_i)$ such that N_i points outward. PIck a partition of unity subordinate to $\{U_i\}$, and let $N = \sum p_i U_i$.
Now, if M is orientable, then there is a non-vanishing top form μ. Define

$$\nu = (\iota(N)\mu)\Big|_{\partial D}.$$

Claim II: $\nu_x \neq 0$ for all $x \in \partial D$.
In adapted coordinates,

$$\mu = f(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n, \quad f \neq 0$$

$$N = N_1 \frac{\partial}{\partial x_1} + \dots + N_n \frac{\partial}{\partial x_n}$$

Then

$$\begin{aligned} \iota(N)\mu\Big|_{x_1=0} &= N_1 f dx_2 \wedge \dots \wedge dx_n \\ &= \sum_{j=2}^{n} (-1)^j f N_j dx_1 \wedge \dots \wedge \hat{dx_j} \wedge \dots \wedge dx_n\Big|_{x_1=0} \\ &= (N_1 f)\Big|_{x_1=0} dx_2 \wedge \dots \wedge dx_n. \end{aligned}$$

□

Now we're ready to discuss Stokes's Theorem.

Theorem 7.12. Stokes's Theorem
Let M be an oriented n-dimensional manifold, $D \in M$ a domain, and $\omega \in \Omega_C^{n-1}(M)$. Then

$$\int_{int(D)} d\omega = \int_{\partial D} \omega.$$

Proof. First, consider the case that $M = \mathbb{R}^n$ and $D = \{x : x_1 \leq 0\}$. Now, $\mu = dx_1 \wedge \dots \wedge dx_n$ is an orientation on M. Let $N = \frac{\partial}{\partial x_1}$ so that $\iota(N)\mu\Big|_{\partial D} = dx_2 \wedge \dots \wedge dx_n$. Let $\omega \in \Omega_c^{n-1}(\mathbb{R}^n)$. Then

$$\omega = \sum_j (-1)^{j-1} f_j dx_1 \wedge \dots \wedge \hat{dx_j} \wedge \dots \wedge dx_n.$$

So

$$d\omega = \sum_j (-1)^{j-1} \frac{\partial f}{\partial x_j} dx_j \wedge dx_1 \wedge \ldots \wedge \hat{dx_j} \wedge \ldots \wedge dx_n$$

$$= \sum_j \frac{\partial f_j}{\partial x_j} dx_1 \wedge \ldots \wedge dx_n.$$

Now,

$$\int_D d\omega = \sum_j \int_{\{x_1 \leq 0\}} \frac{\partial f}{\partial x_j} dx_1 \wedge \ldots \wedge dx_n.$$

Since the support of f_j is compact for all j, there is an $R > 0$ such that $\mathrm{supp}(f_j) \subset \{x \in \mathbb{R}^n : -R \leq x_j \leq R\}$. Then if $j > 1$,

$$\begin{aligned}
\int_{\{x_1<0\}} \frac{\partial f}{\partial x_j} dx_j &= \int \int_{\mathbb{R}} \frac{\partial f}{\partial x_j} dx_j \\
&= \int_{\mathbb{R}} (\int_{-R}^{R} \frac{\partial f}{\partial x_j} dx_j) \\
&= 0.
\end{aligned}$$

For $j = 1$, we have

$$\begin{aligned}
\int_{\{x_1<0\}} \frac{\partial f_1}{\partial x_1} dx_1 &= \int_{-\infty}^{0} \frac{\partial f_1}{\partial x_1} dx_1) dx_2 \ldots dx_n \\
&= f_1(0, x_2, \ldots, x_n).
\end{aligned}$$

Combining these results, we have

$$\int_D d\omega = \int_{\{0\} \times \mathbb{R}^{n-1}} f_1(0, x_2, \ldots, x_n) = \int_{\partial D} \omega\Big|_{\partial D}.$$

Now let's consider the general case. So let M be a manifold, and suppose that $\omega \in \Omega_c^{n-1}(M)$. Suppose further that there is a coordinate chart (U, ϕ) adapted to D with $\mathrm{supp}(d\omega) \subset U$ and

$$\begin{aligned}
\int_{\mathrm{int}(D)} d\omega &= \int_{\mathrm{int}(D) \cap U} d\omega \\
&= \int_{\phi(\mathrm{int}(D))} \phi^*(d\omega) \\
&= \int_{\{x_1<0\}} d(\phi^* \omega)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\{0\}\times\mathbb{R}^{n-1}} \phi^*\omega\Big|_{\{0\}\times\mathbb{R}^{n-1}} \\
&= \int_{U\cap\partial D} \phi^*\omega\Big|_{\phi(U\cap\partial D)} \\
&= \int_{U\cap\partial D} \omega\Big|_{U\cap\partial D} \\
&= \int_{\partial D} \omega\Big|_{\partial D}.
\end{aligned}
$$

Finally, given ω as above, cover D by adapted charts $\{(U_\alpha, \phi_\alpha)\}$. We may assume that $M = \cup_\alpha U_\alpha$. Let $\{p_\alpha\}$ be a partition of unity subordinate to $\{U_\alpha\}$. Then $\sum_\alpha p_\alpha\omega = \omega$ and $\mathrm{supp}(p_\alpha\omega) \subset U_\alpha$. So

$$
\begin{aligned}
\int_{\mathrm{int}(D)} d\omega &= \int_{\mathrm{int}(D)} \sum_\alpha d(p_\alpha\omega) \\
&= \sum_\alpha \int_{\mathrm{int}(D)} d(p_\alpha\omega) \\
&= \sum_\alpha \int_{\partial D} p_\alpha\omega\Big|_{\partial D} \\
&= \int_{\partial D} \sum_\alpha p_\alpha\omega\Big|_{\partial D} \\
&= \int_{\partial D} \omega\Big|_{\partial D}.
\end{aligned}
$$

□

Stokes's Theorem generalizes to boundaries that are smooth a.e.
As promised, we now compute the de Rham groups of S^1.

Example 7.13. de Rham cohomology of S^1
Since S^1 is connected, we know already that $H^1(S^1) = \mathbb{R}$. Furthermore, since S^1 has dimension 1, we know that $H^k(S^1) = \{0\}$ for $k > 1$. So the only question is what happens when $k = 1$.
Now, $d\theta$ is not exact, for if it were, then its integral over S^1 would be 0 (Stoke's Theorem) rather than 2π. Also, note that all 1-forms on S^1 are closed. We claim that if α is a 1-form, then $\alpha - kd\theta$ is exact for some $k \in \mathbb{R}$. For let $\alpha = f(\theta)d\theta$, and set

$$k = \frac{1}{2\pi}\int_{S^1} \alpha.$$

Let

$$g(\theta) = \int_0^\theta (f(\theta) - k)d\theta.$$

Note that since g is 2π-periodic, it's a well-defined C^∞ function on S^1, and $dg = (f(\theta) - k)d\theta = \alpha - kd\theta$. Thus, every 1-form on S^1 differs from a real multiple of $d\theta$ by an exact form. Consequently, $H^1(S^1) \simeq \mathbb{R}$. $\square$

Exercise 7.4.
Let M be an n-dimensional compact oriented manifold, $D \subset M$ a domain with smooth boundary, $f \in C^\infty(M)$, and $\omega \in \Omega^{n-1}(M)$. Show that

$$\int_D f\ d\omega = \int_{\partial D} f\omega - \int_D df \wedge \omega.$$

Exercise 7.5.
Let M be an n-dimensional oriented manifold and $\mu \in \Omega^n(M)$ a nowhere vanishing form. Show that for any vector field X on M, the Lie derivative $L_X\mu$ satisfies n-form

$$L_X\mu = f\mu$$

for some $f \in C^\infty(M)$. We define the **divergence** of X with respect to μ to be this function f and denote it by $\mathrm{div}_\mu(X)$. Thus, $L_X\mu = \mathrm{div}_\mu(X)$. Show that for $M = \mathbb{R}^n$, $\mu = dx_1 \wedge \ldots \wedge dx_n$

$$\mathrm{div}_\mu(\sum_i v^i \frac{\partial}{\partial x_i}) = \sum_i \frac{\partial v^i}{\partial x_i}.$$

Show that if $D \subset M$ is a domain with smooth boundary then

$$\int_D \mathrm{div}_\mu(X) = \int_{\partial D} \iota(X)\mu$$

for any vector field X with compact support.

Exercise 7.6.
Let M be a compact, oriented n-manifold without boundary. Show that $H^n(M) \neq \{0\}$.

Exercise 7.7.
Compute the integral of $x\ dy - y\ dx$ over ∂D, where D is the unit disk in $\mathbb{R}^2$.

Exercise 7.8.
Evaluate $\int_S \omega\Big|_S$ where S is the helicoid in $\mathbb{R}^3$ parameterized by $\phi(s,t) = (s\cos t, s\sin t, t)$, $0 < s < 1$, $0 < t < 4\pi$, and $\omega = z\ dx \wedge dy + 3\ dz \wedge dx - x\ dy \wedge dz$. Use the orientation of S defined by ϕ.

Exercise 7.9.
Supose that ω is any 1-form on S^1 such that $\int_{S^1} \omega = 1$. Let $f(\theta) = \sin 5\theta + 3\theta$. Compute $\int_{S^1} f^*\omega$.

Exercise 7.10.
Let $S^2 \subset \mathbb{R}^3$ be the unit sphere. Let n be the north pole and

$$s : S^2 \smallsetminus \{n\} \to \mathbb{R}^2$$

be the stereographic projection. Let $\omega_0 = dx \wedge dy$ be the standard area form on $\mathbb{R}^2$.
(a) Choose a coordinate system on S^2 and use it to compute $s^*\omega_0$.
Hint: With a clever choice of coordinates, computation is not necessary!
(b) Compute the integral of $s^*\omega_0$ over the lower hemisphere.

8 Riemannian Geometry

8.1 Connections on Vector Bundles

If M is a manifold, then comparing tangent vectors from different fibers will tell us something about the curvature of our manifold. Our initial goal in defining connections is **parallel transport**: That is, if $\gamma : [a, b] \to M$ is a curve on M, we want linear maps $P_{\gamma(t)} : E_{\gamma(a)} \to E_{\gamma(b)}$ such that $P_{\gamma(t)}$ depends smoothly on t. This will allow us to compare vectors from different fibers. The key tool for us to define parallel transport is connections on manifolds, but specifically, we talk about connections on vector bundles. While vector fields define differential operators that act $C^\infty(M)$ functions, connections define a sort of derivative on vector fields (or sections of the appropriate vector bundle).

Definition 8.1. Connection on a Vector Bundle
Let $\pi : E \to M$ be a vector bundle. A **connection** (or **covariant derivative**) is an $\mathbb{R}$-bilinear map $\nabla : \Gamma(TM) \times \Gamma(E) \to \Gamma(E)$, denoted by $(X, s) \mapsto \nabla_X s$, such that for all $f \in C^\infty(M)$, all $X \in \Gamma(TM)$, and all $s \in \Gamma(E)$,
(1) $\nabla_{fX} s = f \nabla_X s$
(2) $\nabla_X(fs) = X(f) \cdot s + f \nabla_X s$.

Let's see an example of this.

Example 8.2.
Suppose that $\pi : E \to M$ is a trivial bundle of rank k. Then there exist global sections $\{s_1, \ldots, s_k\}$ such that $\{s_j(x)\}$ is a basis for E_x, for all $x \in M$. So for any $s \in \Gamma(E)$, we have $s = \sum_j f_j s_j$, for some C^∞ functions f_j. Let us define

$$\nabla_X(\sum_j f_j s_j) = \sum_j X(f_j) s_j.$$

Let's check that this is indeed a connection on E. First,

$$\nabla_{gX}(\sum_j f_j s_j) = \sum_j g X(f_j) s_j = g \sum_j f_j s_j = g \nabla_X(\sum_j f_j s_j).$$

Next,

$$\begin{aligned}
\nabla_X(g \sum_j f_j s_j) &= \sum_j X(g f_j) s_j \\
&= X(g) \sum_j f_j s_j + g \sum_j X(f_j) s_j \\
&= X(g)(\sum_j s_j f_j) + g \nabla_X(\sum_j f_j s_j).
\end{aligned}$$

□

Proposition 8.3.
Let ∇ be a connection on a vector bundle $\pi : E \to M$. Then ∇ is local: That is, for any open set U and any vector fields X and Y, and for any sections s and s' of E such that $X\Big|_U = Y\Big|_U$ and $s\Big|_U = s'\Big|_U$, we have $(\nabla_X s)\Big|_U = (\nabla_Y s')\Big|_U$.

Proof. Since ∇ is bilinear, it is enough to show:
(a) If $X\Big|_U = 0$, then $\nabla_X s\Big|_U = 0$ for any $s \in \Gamma(E)$.
(b) If $s\Big|_U = 0$, then $\nabla_X s\Big|_U = 0$ for any $X \in \Gamma(TM)$.
Pick $x_0 \in U$. Then there is an open set $V \subset U$, $x_0 \in V$, and a function $\rho \in C^\infty_C(U)$ such that $\rho \equiv 1$ on V. Then if $X\Big|U = 0$, $\rho X \equiv 0$, and hence for any section s of E,

$$0 = (\nabla_{\rho X} s)(x_0) = \rho(x_0)(\nabla_X s)(x_0) = (\nabla_X s)(x_0).$$

As $x_0 \in U$ is arbitrary, (a) follows.
Now, if $s\Big|_U = 0$, $\rho s = 0$ on M, which in turn implies that

$$\begin{aligned} 0 &= (\nabla_X \rho s)(x_0) \\ &= (X(\rho)s + \rho \nabla_X s)(x_0) \\ &= 0 + \rho(x_0)(\nabla_X s)(x_0) \\ &= (\nabla_X s)(x_0). \end{aligned}$$

□

Thus, if ∇ is a connection on E, then ∇ induces a connection ∇^U on $E\Big|_U$. Recall that given $U \subset M$ and $s \in \Gamma(E\Big|_U)$, for any $x_0 \in U$, there is $\tilde{s} \in \Gamma(E)$ such that $\tilde{s}(x) = s(x)$ for x near x_0. In particular, for any $X \in \Gamma(TU)$, there is $\tilde{X} \in \Gamma(TM)$ such that $X(x) = \tilde{X}(x)$ for any x near x_0. By the previous proposition,

$$(\nabla^U_X s)(x) = (\nabla_{\tilde{X}} \tilde{s})(x)$$

is a well-defined expression for all x sufficiently close to x_0.

Definition 8.4. Christoffel Symbols
Let U be an open subset of M on which we have a coordinate chart $(U, (x_1, \ldots, x_n))$ and on which E is trivial. Let $\{s_\alpha\}$ be a frame of E on U. Then

$$\nabla_{\frac{\partial}{\partial x_i}} s_\alpha = \sum_\beta \Gamma^\beta_{i\alpha} s_\beta$$

for some functions $\Gamma^\beta_{i\alpha} \in C^\infty(U)$. These functions are the **Christoffel symbols of the connection ∇ relative to $(x_1, \ldots, x_n)$ and $\{s_\alpha\}$.**

A connection is uniquely determined by its Christoffel symbols, a fact which we can deduce directly from the definitions.

Proposition 8.5.
Let ∇ be a connection on $\pi : E \to M$. If $X(x) = 0$, then $(\nabla_X s)(x) = 0$. Thus, $(\nabla_X s)(x)$ depends only upon $X(x)$.

Proof. Choose a coordinate chart $(U, (x_1, \ldots, x_n))$ s.t. $x_0 \in U$ and $E\Big|_U$ is trivial. Pick a local frame $\{s_j\}$ of $E\Big|_U$. At this point, it is convenient to use Einstein summation convention, where we omit summations but sum whenever indices are repeated; i.e., $a_i b_i$ means $\sum_i a_i b_i$. Now,

$$\begin{aligned}
\nabla_{X^i \frac{\partial}{\partial x_i}}(f_j s_j) &= X^i \nabla_{\frac{\partial}{\partial x_i}}(f_j s_j) \\
&= X^i \frac{\partial f_j}{\partial x_i} s_j + X_i f_j \nabla_{\frac{\partial}{\partial x_i}} s_j \\
&= X^i (\frac{\partial f_j}{\partial x_i} s_j + f_j \Gamma^k_{ij} s_k) \\
&= X^i (\frac{\partial f_j}{\partial x_i} + f_\alpha \Gamma^j_{i\alpha}) s_j.
\end{aligned}$$

If $X(x_0) = 0$, $X^i(x_0) = 0$ for all i, which means that $(\nabla_X(f_j s_j))(x_0) = 0$. $\square$

Here's a useful expression which we've obtained as a result of the preceding proof:

Corollary 8.5.1.

$$\nabla_{\sum_i X^i \frac{\partial}{\partial x_i}} (\sum_j f_j s_j) = \sum_{i,j} X^i (\frac{\partial f_j}{\partial x_i} + \sum_\alpha f_j \Gamma^j_{i\alpha}) s_j.$$

The next proposition will complete our "basic material" on connections. It says that connections always exist, and as the reader should expect by now, we'll use the previous example from this section to construct connections on sufficiently small open sets (where the bundle is trivial), and then we'll add each of the connections together using a partition of unity.

Proposition 8.6.
Let $\pi : E \to M$ be a vector bundle. Then there is a connection on E.

Proof. Choose $\{U_\alpha\}$ on M such that $E\Big|_{U_\alpha}$ is trivial. Let ∇^α be a connection on $E\Big|_{U_\alpha}$, as outlined in the previous example. Furthermore, let $\{p_\alpha\}$ be a partition of unity subordinate to $\{U_\alpha\}$, and define

$$\nabla_X s = \sum_\alpha p_\alpha(\nabla_X^\alpha(s\Big|_{U_\alpha})).$$

That this is a connection now follows from the details of the previous example, and the details are left to the reader. □

8.2 Parallel Transport

Recall the goal of parallel transport: Let $\pi : E \to M$ be a vector bundle. Then given a curve $\gamma : [a, b] \to M$, we want linear maps

$$P_{\gamma(t)} : E_{\gamma(a)} \to E_{\gamma(t)}$$

which depend smoothly on t.
How can we construct these maps? First note that if for every $v \in E_{\gamma(a)}$, we can find a global section $s^v \in \Gamma(E)$ such that $s^v(\gamma(a)) = v$ and $(\nabla_{\dot{\gamma}(t)} s^v)(\gamma(t)) = 0$, then we can define $P_{\gamma(t)}(v) = s^v(\gamma(t))$. In some sense, this says that we can find a section which is constant along $\dot{\gamma(t)}$. Of course, global sections do not exist in general; however, we can still parallel transport locally, as we will see.

Example 8.7.
Let $M = \mathbb{R}^n$, $E = TM$, $\Gamma(E) = C^\infty(M, \mathbb{R}^n)$. Let

$$\nabla_X Y = (X(Y_1), \dots, X(Y_n)).$$

Then we need

$$0 = \nabla_{\dot{\gamma}(t)} Y = (\frac{d}{dt} Y_1(\gamma(t)), \dots, \frac{d}{dt} Y_n(\gamma(t))),$$

which says that Y is constant along gamma. □

To define local parallel transport, we need the notion of the covariant derivative along a curve. We now define the appropriate concepts.

Definition 8.8.
Let $\pi : E \to M$ be a vector bundle, $\gamma : [a, b] \to M$ a curve. $\sigma : [a, b] \to E$ is a **section of** E **along** γ $\sigma(t) \in E_{\gamma(t)}$ for all $t \in [a, b]$.

Note that if $s \in \Gamma(E)$, $s \circ \gamma$ is a section of E along γ. We denote the collection of sections along γ by $\Gamma(\gamma^*(E))$.

Pull-back Bundles

Here's another way to consider sections along a curve γ. Suppose $f : N \to M$ is smooth and that $\pi : E \to M$ is a vector bundle. Define

$$f^*E = \{(n, e) \in N \times E : f(n) = \pi(e)\}.$$

Exercise 8.1.
1. f^*E is a submanifold of $N \times E$.
Hint: Surjectivity of $df + d\pi$ and a previous homework exercise on transversality.
2. f^*E is a vector bundle.
Hint: Local sections of f^*E are given by $\{s_i \circ f\}$, where $\{s_i\}$ is a local frame for E.

What's the big deal? Well, if σ is a section of a vector bundle E along a curve γ, then σ is a section of $\gamma^*E = \{(t, e) : \gamma(t) = \pi(e)\}$. Conversely, a section of $\gamma^*E \to [a, b]$ is of the form $t \mapsto (t, \sigma(t))$, where $\pi(\sigma(t)) = \gamma(t)$. Thus, sections of γ^*E *are* sections of E along γ!
As a side note, if $s \in \Gamma(E)$, then $s \circ \gamma \in \Gamma(\gamma^*E)$. However, this assignment $\Gamma(E) \to \Gamma(\gamma^*E)$ need not be surjective (consider any curve that intersects itself!)

Covariant Derivatives

Here's the piece of machinery that will allow us to write a differential equation which more or less defines parallel transport locally.

Definition 8.9. Covariant Derivative
A covariant derivative $\frac{\nabla}{dt}$ along γ is an $\mathbb{R}$-linear map $\frac{\nabla}{dt} : \Gamma(\gamma^*(E)) \to \Gamma(\gamma^*(E))$ given by $\sigma \mapsto \frac{\nabla}{dt}\sigma$ such that for all $f \in C^\infty([a, b])$,

$$\frac{\nabla}{dt}(f\sigma) = \frac{df}{dt}\sigma + f\frac{\nabla}{dt}\sigma.$$

Proposition 8.10.
Given a connection ∇ on $\pi : E \to M$ and a curve $\gamma : [a,b] \to M$, there is a unique covariant derivative $\frac{\nabla}{dt} : \Gamma(\gamma^*(E)) \to \Gamma(\gamma^*(E))$ along γ such that

$$\frac{\nabla}{dt}(s \circ \gamma)(t) = (\nabla_{\dot\gamma(t)} s)(\gamma(t)). \tag{2}$$

Proof. **Uniqueness**: Suppose that $U \subset M$, $\frac{\nabla}{dt}$ satsfies (2), and $E\Big|_U$ is trivial. Pick a frame $\{s_j\}$ of $E\Big|_U$, and let $\{s_j^*\}$ be the dual frame. Further let $I = \gamma^{-1}(U)$. If $s \in \Gamma(E)$, then $s\Big|_U = \sum_j \langle s_j^*, s\rangle s_j$, which in turn implies that if $\sigma \in \Gamma(\gamma^* E\Big|_I)$, then

$$\begin{aligned} \sigma &= \sum_j \langle s_j^* \circ \gamma, \sigma\rangle (s_j \circ \gamma) \Rightarrow \\ \frac{\nabla}{dt}\sigma &= \frac{\nabla}{dt}(\sum_j \sigma_j (s_j \circ \gamma)) \end{aligned}$$

Thus,

$$\sum_j \frac{d\sigma_j}{dt} s_j \circ \gamma + \sum_j \sigma_j (\nabla_{\dot\gamma} s_j) \circ \gamma, \tag{3}$$

where $\sigma_j \in C^\infty(I)$. Since all of these parameters are uniquely determined, $\frac{\nabla}{dt}$ must be unique.
Existence: Cover $\gamma([a,b])$ with sets U_j such that $E\Big|_{U_j}$ is trivial. It's enough to construct $\frac{\nabla}{dt}$ on each $\Gamma(\gamma^* E\Big|_{\gamma^{-1}(U_j)})$. By uniqueness, we then get $\frac{\nabla}{dt}$ on all of $\Gamma(\gamma^* E)$. Pick a frame $\{s_k^j\}$ on $E\Big|_{U_j}$ and define $\frac{\nabla}{dt}$ on $\gamma^* E$ by (3). This is a covariant derivative satisfying (2). $\square$

To define parallel transport along a curve $\gamma : [a,b] \to M$, we want, for every $v \in E_{\gamma(a)}$, a section $\sigma^v \in \Gamma(\gamma^*(E))$ such that $\sigma^v(a) = v$ and $\frac{\nabla}{dt}\sigma^v = 0$. We also want the map $v \mapsto \sigma^v$ to be linear. Thus, we can define parallel transport on γ if we compute (3) in coordinates. Suppose that $\dot\gamma = \sum_i (\frac{d}{dt}\gamma^i)\frac{\partial}{\partial x_i}$, with $\gamma^i = x_i \circ \gamma$. Then if $\sigma = \sum \sigma_\alpha (s_\alpha \circ \gamma)$ in a coordinate chart that trivializes E, $0 = \frac{\nabla}{dt}\sigma$ amounts to

$$0 = \frac{\nabla}{dt}\sigma$$

$$
\begin{aligned}
&= \sum_j \frac{d\sigma_j}{dt}(s_j \circ \gamma) + \sum_{i,j} \sigma_j \dot{\gamma}^i (\nabla_{\frac{\partial}{\partial x_i}} s_j) \circ \gamma \\
&\Rightarrow \\
0 &= \sum_j \frac{d\sigma_j}{dt}(s_j \circ \gamma) + \sum_{i,j,k} \sigma_j \dot{\gamma}^i (\Gamma_{ij}^k s_j) \circ \gamma,
\end{aligned}
$$

and changing indices, we see that

$$\frac{d\sigma_\alpha}{dt} = -\sum_{i,\delta} \sigma_\delta \dot{\gamma}_i (\Gamma_{i\delta}^\alpha \circ \gamma).$$

In other words, $\frac{\nabla}{dt}\sigma = 0$ amounts to

$$\frac{d}{dt}\begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_k \end{pmatrix} = B(t)\frac{d}{dt}\begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_k \end{pmatrix},$$

where

$$B_{jk} = -\sum_i \dot{\gamma}^i (\Gamma_{ik}^j \circ \gamma).$$

These last two equations are two equivalent versions of the **parallel transport equation**. We can turn this into a system of ODE's and (in theory) solve to get what the σ_α terms should be. Once we have the solution on a coordinate chart, we can then piece solutions from different coordinate charts together to get the result.

The relevant theorem is the following one:

Theorem 8.11.
Suppose that $B : [c,d] \to M_k(\mathbb{R})$ is a smooth curve in the space of $k \times k$ real matrices. Then there is a map $R : [c,d] \to M_k(\mathbb{R})$ such that $\sigma(t) = R(t)\sigma^0$ is a solution with inital conditions $\sigma(c) = \gamma^0$.

8.3 Riemannian Manifolds

Let $x \in M$, where M is some manifold. Then T_xM is a real vector space, and as such, we can define an inner product on it in a number of ways. This idea leads to the following concept.

Definition 8.12. Riemannian Metric
A **Riemannian metric** g on a manifold M assigns smoothly to each $x \in M$ an inner product g_x on T_xM.

Here, "smoothly" means that g is a section of the vector bundle of symmetric bilinear maps from $TM \times TM \to \mathbb{R}$ and that the functions $g_{ij} = g(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})$ are smooth on the coordinate chart $(U, (x_1, \ldots, x_n))$.

Theorem 8.13.
Any manifold has a Riemannian metric.

Proof. Let $\{(U_i, \phi_i = (x_1^i, \ldots, x_n^i))\}$ be a collection of coordinate charts that cover M. Let $\{p_i\}$ be a partition of unity subordinate to this collection, and define

$$g = \sum_i p_i(\sum_j dx_j^i \otimes dx_j^i).$$

Then g is a Riemannian metric. Symmetry and positive definiteness are easy to verify, and smoothness follows from the smoothness of the $p_i's$ and the fact that $\sum_j dx_j^i \otimes dx_j^i$ is a nonvarying bilinear map on TU_i. $\square$

Definition 8.14. Riemannian Manifold
A **Riemannian manifold** is a manifold M together with a Riemannian metric g.

It turns out that on a Riemannian manifold, there is a special type of connection. This connection is called the **Levi-Civita connection** or sometimes the **Riemannian connection**.

Theorem 8.15. Levi-Civita Connection
On every Riemannian manifold (M, g) there is a unique connection $\nabla : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$ which is
(1) Torsion-free : $\nabla_X Y - \nabla_Y X = [X, Y]$
(2) Compatible with g : $X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$

Proof. **Step I: Uniqueness**
Suppose that ∇ exists. Then for any $X, Y, Z \in \Gamma(TM)$,

$$\begin{aligned}
& 2g(Z, \nabla_Y X) + g([X, Y], Y) + g([Y, Z], X) + g([X, Y], Z) \\
= \; & g(Z, \nabla_Y X) + g(Z, \nabla_Y X) + g(\nabla_X Z, Y - \nabla_Z X, Y) \\
& + g(\nabla_Y Z - \nabla_Z Y, X) + g(\nabla_X Y - \nabla_Y X, Z) \\
= \; & (g(Z, \nabla_Z X) + g(\nabla_Y Z, X)) + (g(\nabla_X Z, Y) + g(\nabla_Y Z, X)) \\
& - (g(\nabla_Z X, Y) + g(X, \nabla_Z Y)) \\
= \; & Y(g(Z, X)) + X(g(Z, Y)) - Z(g(X, Y)).
\end{aligned}$$

Thus, we have

$$\begin{aligned} g(Z, \nabla_Y X) &= \frac{1}{2}(X(g(Y,Z)) + Y(g(X,Z)) - Z(g(X,Y)) - g([X,Z],Y) \\ &\quad - g([Y,Z],X) - g([X,Y],Z)). \end{aligned}$$

Since Z is arbitrary and g is nondegenerate, the formula above uniquely determines $\nabla_X Y$.

Step II : Existence : By uniqueness, it will be enough to construct ∇ in a coordinate chart. To do this, we'll compute Christoffel symbols for ∇ in terms of $g_{ij} = g(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})$ and $\frac{\partial}{\partial x_j} g_{ij}$ and then define ∇ in terms of these Christoffel symbols. Finally, we will verify that the connection defined is the Levi-Civita connection.

Before we begin, let us introduce two conventions. First, let ∂_i denote $\frac{\partial}{\partial x_i}$. Next, we will use the Einstein summation convention (we omit $\sum$ symbols, keeping in mind that we sum over repeated indices). For example, by $X_i \partial_i$, we mean $\sum_i X_i \partial_i$.

Since $[\partial_i, \partial_j] = 0$ always, we see that if ∇ is a Levi-Civita connection, then by the equation for $g(Z, \nabla_X Y)$ above, we see that

$$g(\partial_k, \nabla_{\partial_i} \partial_j) = \frac{1}{2}(\partial_j g_{ik} + \partial_i g_{jk} - \partial_k g_{ji}).$$

On the other hand,

$$\nabla_{\partial_i} \partial_j = \Gamma^l_{ij} \partial_l,$$

so that

$$g(\partial_k, \nabla_{\partial_i} \partial_j) = \Gamma^l_{ij} g(\partial_k, \partial_l) = \Gamma^l_{ij} g_{kl}.$$

As g_{ij} is positive definite, it is nondegerate. Let $(g^{kl}) = (g_{ij})^{-1}$. Hence,

$$g^{sk} g_{kl} \Gamma^l_{ij} = \frac{1}{2} g^{sk}(\partial_j g_{ik} + \partial_i g_{jk} - \partial_k g_{ij}).$$

That is,

$$\Gamma^s_{ij} = \frac{1}{2} g^{sk}(\partial_j g_{ik} + \partial_i g_{jk} - \partial_k g_{ij}).$$

Define ∇ in coordinates to be the connection whose Christoffel symbols given by the previous equation. We need to check that this is the Levi-Civita connection.

Since $\Gamma^k_{ij} = \Gamma^k_{ji}$,

$$\nabla_{\partial_i} \partial_j - \nabla_{\partial_j} \partial_i = \Gamma^k_{ij} \partial_k - \Gamma^k_{ji} \partial_k = 0.$$

Thus, for two vector fields $X = X^i\partial_i$ and $Y = Y^i\partial_i$, we have

$$\begin{aligned}
\nabla_X Y - \nabla_Y X &= \nabla_{X^i\partial_i}(Y^j p_j) \\
&= X^i(\partial_i Y^j)\partial_j + X^iY^j\nabla_{\partial_i}\partial_j - Y^j(\partial_j X^i)\partial_i - Y^jX^i\nabla_{\partial_j}\partial_i \\
&= X^i(\partial_i Y^j)\partial_j - Y^j(\partial_j X^i)\partial_i \\
&= [X^i\partial_i, Y^j\partial_j].
\end{aligned}$$

Thus, ∇ is torsion-free. Compatibility with g is a somewhat longer computation. First, note that

$$\begin{aligned}
g(\nabla_{\partial_i}\partial_j, \partial_k) + g(\partial_j, \nabla_{\partial_i}\partial_k) &= g(\Gamma^l_{ij}\partial_l, \partial_k) + g(\partial_j, \Gamma^m_{ik}\partial_m) \\
&= \Gamma^l_{ij}g_{lk} + \Gamma^m_{ik}g_{jm} \\
&= \partial_i g_{jk},
\end{aligned}$$

where the last equality follows from our definition of the Christoffel symbols for ∇. Thus, we have for vector fields X, Y, Z,

$$\begin{aligned}
(X^j\partial_j)g(Y^i\partial_i, Z^k\partial_k) &= X^j\partial_j(Y^iZ^kg_{ik}) \\
&= X^j(\partial_j Y^i)Z^kg_{ik} + X^jY^i(\partial_j Z^k)g_{ik} + X^iY^jZ^k(\partial_j g_{ik}) \\
&= g(X^j(\partial_j Y^i)\partial_i, Z^k\partial^k) + g(Y^i\partial_i, X^j(\partial_j Z^k)\partial^k) \\
&\quad + X^jY^iZ^k(g(\nabla_{\partial_j}\partial_i, \partial_k) + g(\partial_i, \nabla_{\partial_j}\partial_k)) \\
&= g((X^j\partial_j)Y^i\partial_i, Z^k\partial^k) + g(Y^i\nabla_{X^j\partial_j}\partial_i, Z^k\partial^k) \\
&\quad + g(Y^i\partial_i, (X^j\partial_j)Z^k\partial_k) + g(Y^i\partial_i, Z^k\nabla_{X^j\partial_j}\partial_k) \\
&= g(\nabla_{X^j\partial_j}(Y^i\partial_i), Z^k\partial^k) + g(Y^i\partial^i, \nabla_{X^j\partial_j}(Z^k\partial_k)).
\end{aligned}$$

That is, ∇ is compatible with g, and it follows that ∇ is the Levi-Civita connection with respect to g. $\square$

Example 8.16.
Let Y be a vector field on $\mathbb{R}^3$. Define $D_X Y = \sum X(Y^i)\frac{\partial}{\partial x_i}$. Then D is the Levi-Civita connection on $\mathbb{R}^3$ (with respect to the standard inner product on $\mathbb{R}^3$. We leave this check to the reader. We also let the reader verify that the Levi-Civita connection on $\mathbb{R}^n$ is also characterized by the fact that all of its Christoffel symbols are zero. $\square$

Exercise 8.2.
Suppose that $\gamma : \mathbb{R} \to \mathbb{R}^n$ is a curve. Show that $\frac{D}{dt}\dot{\gamma} = \ddot{\gamma}$.

Exercise 8.3.
Suppose that (M, g) is a Riemannian manifold. Then g induces a bundle isomorphism $g^{\#} : TM \to T^*M$ by

$$(x, v) \mapsto g_x(v, \cdot).$$

A smooth map $F : (M, g_M) \to (N, g_N)$ is an **isometry** if $F^* g^N = g^M$. Here,

$$F^* g_x^M(v, w) = g_{F(x)}^N(dF_x(v), dF_x(w)).$$

Exercise 8.4.
Show that isometries preserve the Riemannian connection. That is, if $F : (M, g^M) \to (N, g^N)$ is an isometry and ∇^M and ∇^N are the respective Levi-Civita connections, then

$$F^*(\nabla_X^M Y) = \nabla_{dF(X)}^N dF(Y).$$

Induced Connections

We now consider an important example of Levi-Civita connections; namely, if we have an embedded submanifold of a larger manifold, how can we define the Levi-Civita connection on this submanifold? In other words, if (M, g_M) is a Riemannian manifold and N is a closed submanifold, then the inclusion induces a Riemannian metric on N by $g^N = \iota^* g^M$. That is, if $x \in N$, we can define g^N on a pair of vectors (v, w) by

$$g_x^N(v, w) = g_{\iota(x)}^M(d\iota(v), d\iota(w)).$$

It is a fact, left to the reader to pursue, that g^N is a Riemannian metric on N.

Exercise 8.5.
If $\iota : N \to M$ is an embedding and g^M is a Riemannian metric on M, then $\iota^* g^M$ is a Riemannian metric on N.

Armed with these definitions, we can now consider the relationship between the two corresponding Levi-Civita connections ∇^M and ∇^N. (The connection on N sometimes is called an **induced connection**). To begin, we need a technical result.

Lemma 8.17.
Let N be a closed submanifold of a manifold M. Let X and Y be vector fields on N, and let W and Z be extensions of them to M. Then for any point $x \in N$, $\nabla_W^M Z$ depends only on X and Y and not their extensions.

Proof. $\nabla^M_W Z$ depends only on $W(x) = X(x)$ and on the values of Z along the integral curve γ through x. As W is tangent to N, $\gamma \subset N$ and $Z(\gamma(t)) = Y(\gamma(t))$, for all t. $\square$

It is *not* true, however, that $(\nabla^M_W Z)(x) \in T_xN$. To see this, consider the following example.

Example 8.18. Let $W = Z = x_2\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_2}$, a vector field on $\mathbb{R}^2$ that corresponds to a circular flow around the origin. If D represents the Levi-Civita connection on $\mathbb{R}^2$, we have

$$D_W Z = (W)(x_2)\frac{\partial}{\partial x_1} + (W)(-x_1)\frac{\partial}{\partial x_2} = -x_1\frac{\partial}{\partial x_1} - x_2\frac{\partial}{\partial x_2}.$$

Take $M = \mathbb{R}^2$, $N = S^1$. Then $D_W Z$ is orthogonal to S^1 (it is the radial vector field). $\square$

Now, for every $x \in N$, there is an orthogonal projection $\pi_x : T_xM \to T_xN$; that is, there is a vector bundle map $\pi : TM\Big|_N \to TN$. If $(x_1, \ldots, x_n, \ldots, x_m)$ are coordinates for M near $x_0 \in N$ that are adapted to N, we may use the Riemannian metric g to apply Gram-Schmidt to the basis vectors $\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}, \ldots, \frac{\partial}{\partial x_m}\}$ to obtain an orthonormal frame $\{e_1(x), \ldots, e_n(x), \ldots, e_m(x)\}$ on TU. Define

$$\pi_x(v) = \sum_{i=1}^{n} g^M_x(v, e_i(x))e_i(x)$$

and define ∇ on N by

$$(\nabla_X Y)(x) = \pi_x((\nabla^M_{\tilde{X}}\tilde{Y})(x)),$$

where $\tilde{X}$ and $\tilde{Y}$ are any extensions of X and Y.

Proposition 8.19.
$\nabla = \nabla^N$, the Levi-Civita connection on N.

Proof. We show that ∇ preserves the metric; the other two properties are somewhat straightforward and are left to the reader. Thus, we need to check that

$$Z(g^N(X,Y)) = g^N(\nabla_Z X, Y) + g^N(X, \nabla_Z Y)$$

for any vector fields X, Y, Z on N. At any point of N,

$$\begin{aligned}
Z(g^N(X,Y)) &= \tilde{Z}(g^M(\tilde{X},\tilde{Y})) \\
&= g^M(\nabla^M_{\tilde{Z}}\tilde{X},\tilde{Y}) + g^M(\tilde{X},\nabla^M_{\tilde{Z}}\tilde{Y}) \\
&= g^M(\nabla_Z X + (\nabla^M_{\tilde{Z}}\tilde{X} - \nabla_Z X), Y) + g^M(X, \nabla_Z Y + (\nabla^M_{\tilde{Z}}\tilde{Y} - \nabla_Z Y)) \\
&= g^M(\nabla_Z X, Y) + g^M(X, \nabla_Z Y).
\end{aligned}$$

Here we tacitly used the definition of g^N and the fact that $T_xN \subset T_xM$ for all $x \in N$. $\square$

Why are induced connections useful? Suppose, for example, that I want to compute $\frac{\nabla}{dt}\dot{\gamma}$ for some curve in an embedded surface in $\mathbb{R}^3$, I can first compute this quantity in $\mathbb{R}^3$ (which is just $\ddot{\gamma}$) and project the resulting vector onto the tangent plane to my surface. That is, if D is the Levi-Civita connection on $\mathbb{R}^3$, X and Y are vector fields on S, and W and Z are respective extensions to $\mathbb{R}^3$, one can define a connection on S by

$$\nabla_X Y = D_W Z - \langle D_W Z, n\rangle n,$$

where n is the unit normal.

Exercise 8.6.
Consider the mapping $\pi : (u, v) \mapsto (\cos u, \sin u, v)$ which takes $\mathbb{R}^2$ to the cylinder $M = \{(x, y, z) : x^2 + y^2 = 1\}$. Show that π is an **isometry** with respect to the induced metric on M; that is, that $\pi^* g^M = g$, where g is the standard metric on $\mathbb{R}^2$.

8.4 Curvature

Let ∇ be a connection on a vector bundle $E \to M$. For $X, Y \in \Gamma(TM)$, $s \in \Gamma(E)$, the **curvature of** ∇ is defined to be

$$K(X, Y)s = \nabla_X(\nabla_X s) - \nabla_Y(\nabla_X s) - \nabla_{[X,Y]}s.$$

It is evident that this formula defines a multilinear mapping $\Gamma(TM) \times \Gamma(TM) \times \Gamma(E) \to \Gamma(E)$. At first, it is not clear what this definition means. One motivation is as follows. It is a theorem of advanced calculus that the second-order partial derivatives are independent of order. For functions on manifolds and vector fields X and Y, the analogous property does not hold, and in fact, $[X, Y]f = X(Yf) - Y(Xf)$ measures the extent to which $Y(Xf) = X(Yf)$. Since interchangeability of order of differentiation is measured by an "interesting" object $[X, Y]$, one might ask if there is, by analogy, an "interesting" object that measures similar properties for ∇_X and ∇_Y derivatives of a vector field Z on a manifold M. Now, in general, it is *not* the case that $\nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) = 0$; thus, this commutator determines some nonzero vector field on M which may be thought of as an analogue of $[X, Y]$. Curvature is a variant of this expression which also involves noninterchangeability of derivatives of $[X, Y]$.
Curvature is a primary topic of study in Riemannian geometry, and we will

only briefly touch upon it here. Our main focus will be on the curvature of the Levi-Civita connection and its implications to the curvature of surfaces in $\mathbb{R}^3$. First, we derive a few preliminary properties.

Proposition 8.20.
For any $x \in M$, $X, Y \in \Gamma(TM)$ and $s \in \Gamma(E)$, $(K(X,Y)s)(x)$ depends only upon $X(x)$, $Y(x)$, and $s(x)$.

Proof. It suffices to check that K is $C^\infty(M)$-linear, in each of three slots. Accordingly, let $f \in C^\infty(M)$, $X, Y \in \Gamma(TM)$, and $s \in \Gamma(E)$. Then

$$\begin{aligned}
K(X,Y)(fs) &= \nabla_X(\nabla_Y(fs)) - \nabla_Y(\nabla_X(fs)) - \nabla_{[X,Y]}(fs) \\
&= \nabla_X(Y(f)s + f\nabla_Y s) - \nabla_Y(X(f)s - f\nabla_X s) - ([X,Y]f)s - f\nabla_{[X,Y]}s \\
&= X(Y(f))s + Y(f)\nabla_X s + X(f)\nabla_Y s + f\nabla_X(\nabla_Y s) - Y(X(f))s \\
&\quad -X(f)\nabla_Y s - Y(f)\nabla_X s - f\nabla_X(\nabla_Y s) - ([X,Y]f)s - f\nabla_{[X,Y]}s \\
&= fK(X,Y)s
\end{aligned}$$

That $K(fX,Y)s = fK(X,Y)s$ and $K(X,fY)s = f(X,Y)s$ are verified similarly. □

In other words, $K(X,Y)$ is **tensorial**. Note that for each $x \in M$, we have a trilinear map

$$K_x : T_xM \times T_xM \times E_x \to E_x$$

that is skew-symmetric in X and Y. Now we consider a special case, the curvature of the Levi-Civita connection.

Definition 8.21.
If (M,g) is a Riemannian manifold and ∇^g is the correpsonding Levi-Civita connection, then the corresponding curvature R is called the **Riemannian curvature tensor**.

Specifically, we want to consider the case that we have an embedded surface $S \subset \mathbb{R}^3$. For any $x_0 \in S$, we have a neighborhood $U \subset S$ and a vector field $n : U \to \mathbb{R}^3$ such that

$$\begin{aligned}
&(1) \qquad n(x) \perp T_xS \\
&(2) \qquad ||n(x)|| = 1
\end{aligned}$$

Of course, if S is locally the graph of some function $f : \mathbb{R}^2 \to \mathbb{R}$, then $n(x)$ is merely the gradient vector field obtained by

$$(x_0, y_0) \to \Big(\frac{\partial f}{\partial x}(x_0), \frac{\partial f}{\partial y}(y_0), 1\Big).$$

Also, note that $n : U \to S^2$, the two sphere, and that $dn_x : T_xS \to T_{n(x)}S^2 \simeq T_xS$; these tangent spaces may be identified since both are perpendicular to the unit normal. The map n is commonly known as the **Weingarten map**.

Definition 8.22. Gauss Curvature
$\det dn_x = k(x)$ is the **Gauss curvature** of S at x.

One nice fact to know is that $k(x)$ is independent of choice of n; this is part of the content of the next theorem.

Theorem 8.23.
Let $S \subset \mathbb{R}^3$ be an embedded surface. Then for any orthonormal basis e_1, e_2 of T_xS,

$$g_x^S(R_x^S(e_1, e_2)e_1, e_2) = k(x).$$

That is, k is dependent only upon the induced metric g^S.

Let us see a couple of examples of the Gauss curvature.

Example 8.24.
Consider

$$S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0\}.$$

Here the condition that $x_3 = 0$ forces the vector field $n(x)$ to be constant, and so $k(x)$ is 0. $\square$

Example 8.25.

$$S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_2^2 + x_3^2 = R^2\},$$

a cylinder. Here, $n(x)$ is constant in the x_1 direction. Hence, $dn_x(e_1) = 0$, and so $k(x) = 0$.

Example 8.26.

$$S = \{(x_1, x_2, x_3) : x_1^2 + x_2^2 + x_3^2 = R^2\},$$

a sphere. Then $n(x) = \frac{1}{R}$, so that

$$dn = \frac{1}{R} \cdot id.$$

Thus,

$$k(x) = \frac{1}{R^2}.$$

Example 8.27.

$$S = \{(x_1, x_2, x_3) : x_3 = x_1^2 - x_2^2\},$$

a hyperboloid. Then

$$dn_{(0,0,0)} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix},$$

$a > 0$, $b < 0$.

The proof of Gauss's theorem relies strongly on the next two lemmas. In the following, D is the Levi-Civita connection on $\mathbb{R}^3$.

Lemma 8.28.
Let $S \subset \mathbb{R}^3$ be an embedded surface with unit normal n. Let $L = dn$. Then

$$D_{\tilde{X}}\tilde{Y} = \nabla_X Y - \langle L(X), Y\rangle n,$$

where D is the Levi-Civita connection on $\mathbb{R}^3$, ∇ is the connection on S, $X, Y \in \Gamma(TS)$, and $\tilde{X}, \tilde{Y}$ are their extensions.

Proof. At points of S, $\langle n, \tilde{Y}\rangle = 0$ implies that

$$\begin{aligned} 0 &= \tilde{X}\langle \tilde{Y}, n\rangle \\ &= \langle D_{\tilde{X}}\tilde{Y}, n\rangle + \langle \tilde{Y}, D_{\tilde{X}} n\rangle \\ &= \langle D_{\tilde{X}}, n\rangle + \langle \tilde{Y}, L(X)\rangle. \end{aligned}$$

As $D_{\tilde{X}}\tilde{Y} = \nabla_X Y + \langle D_{\tilde{X}}\tilde{Y}, n\rangle n$, we have

$$D_{\tilde{X}}\tilde{Y} = \nabla_X Y - \langle Y, L(X)\rangle n.$$

□

Lemma 8.29.
Let $S \subset \mathbb{R}^3$ be as above, and $X, Y, Z \in \Gamma(TS)$. Then

$$R(X, Y)Z = \langle L(X), Z\rangle L(X) - \langle L(X), Z\rangle L(Y).$$

Proof. Note that

$$\begin{aligned} D_{\tilde{X}}(D_{\tilde{Y}} Z) &= D_{\tilde{X}}(\nabla_Z Y - \langle L(Y), Z\rangle n) \\ &= \nabla_X(\nabla_Y Z) - \langle L(X), \nabla_Y Z\rangle n - X\langle L(Y), Z\rangle n - \langle L(Y), Z\rangle L(X). \end{aligned}$$

Similarly,

$D_{\tilde{Y}}(D_{\tilde{X}}\tilde{Z}) = \nabla_Y(\nabla_X Z) - \langle L(Y), \nabla_X Z\rangle n - Y\langle L(X), Z\rangle n - \langle L(X), Z\rangle L(Y).$

Now,

$$D_{[\tilde{X},\tilde{Y}]}\tilde{Z} = \nabla_{[X,Y]}Z - \langle L([X,Y]), Z\rangle n.$$

Then

$$\begin{aligned} 0 &= k^D(\tilde{X},\tilde{Y})\tilde{Z} \\ &= \nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) - \nabla_{[X,Y]}Z - \langle L(Y), Z\rangle L(X) + \langle L(X), Y\rangle L(Y). \end{aligned}$$

□

Now we can offer a proof of Gauss's theorem.

Proof.

$$\begin{aligned} \langle K(e_1,e_2)e_2, e_1\rangle &= \langle L(e_2), e_2\rangle\langle L(e_1), e_1\rangle - \langle L(e_1), e_2\rangle\langle L(e_2), e_1\rangle \\ &= \det\langle L(e_i), e_j\rangle \\ &= \det L \end{aligned}$$

□

Proposition 8.30.
Let ∇ be the Levi-Civita connection of a Riemannian manifold (M,g). The curvature K is skew-symmetric:

$$g(K(X,Y)Z,W) + g(Z,K(X,Y)W) = 0.$$

Proof. It suffices to show that

$$g(K(X,Y)V,V) = 0.$$

First, note that since ∇ is the Levi-Civita connection,

$$X(g(Z,Z)) = 2g(\nabla_X Z, Z).$$

Then

$$\begin{aligned} 0 &= X(Y(g(Z,Z))) - Y(X(g(Z,Z))) - [X,Y](g(Z,Z)) \\ &= X(2g(\nabla_X(\nabla_Y Z), Z) - Y(2g(\nabla_Y(\nabla_X Z), Z) - 2g(\nabla_{[X,Y]}Z, Z) \\ &= 2g(\nabla_X(\nabla_Y Z), Z) + 2g(\nabla_Y Z, \nabla_X Z) - 2g(\nabla_Y(\nabla_X Z), Z) \\ &\quad -2g(\nabla_X Z, \nabla_Y Z) - 2g(\nabla_{[X,Y]}Z, Z) \\ &= 2g(\nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) - \nabla_{[X,Y]}Z, Z) \\ &= 2g(K(X,Y)Z,Z). \end{aligned}$$

□

In the context of our previous situation, where $S \subset \mathbb{R}^3$ is an embedded surface, the previous lemma says that $K_x(e_1, e_2) : T_xS \to T_xS$ has the form

$$\begin{pmatrix} 0 & K \\ -K & 0 \end{pmatrix}.$$

We may think of K as a 2-form on the manifold whose values are linear transformations on the tangent bundle.

8.5 Geodesics

Definition 8.31. Geodesic
Let ∇ be a connection on M. A curve $\gamma : [a, b] \to M$ is a **geodesic for** ∇if $\frac{\nabla}{dt}\dot{\gamma} = 0$.

Recall that $\dot{\gamma} = \frac{d}{dt}\gamma(t) \in \Gamma(\gamma^*(TM))$. Let's compute the equation $\frac{\nabla}{dt}\dot{\gamma} = 0$ in coordinates and see what we get.

$$\begin{aligned}
0 &= \frac{\nabla}{dt}\dot{\gamma} \\
&= \frac{\nabla}{dt}(\sum_i \dot{\gamma^i}\frac{\partial}{\partial x_i}) \\
&= \sum_i \ddot{\gamma^i}\frac{\partial}{\partial x_i} + \sum_i \dot{\gamma^i}\nabla_{\dot{\gamma}}\frac{\partial}{\partial x_i} \\
&= \sum_k \ddot{\gamma^k}\frac{\partial}{\partial x_k} + \sum_{i,j} \dot{\gamma^i}\dot{\gamma^j}\nabla_{\frac{\partial}{\partial x_j}}\frac{\partial}{\partial x_i} \\
&= \sum_k \ddot{\gamma^k}\frac{\partial}{\partial x_k} + \sum_{i,j} \dot{\gamma^i}\dot{\gamma^j}\Gamma^k_{ij}\frac{\partial}{\partial x_k}.
\end{aligned}$$

That is, γ is a geodesic for ∇ if in coordinates, we have

$$\ddot{\gamma}_k = -\sum_{i,j}\Gamma^k_{ij}\dot{\gamma}_i\dot{\gamma}_j.$$

This is the **geodesic equation**. Geodesics are often defined as "the curves of shortest distance between two points," which is an accurate local description of geodesics which may or may not apply globally.

Exercise 8.7.
γ is a geodesic for the Levi-Civita connection in $\mathbb{R}^3$ if $\ddot{\gamma} = 0$.

Exercise 8.8.
Given a point $x_0 \in M$ and a tangent vector v, there is a geodesic $\gamma : [a, b] \to M$ with $\gamma(0) = x_0$ and $d\gamma(\frac{d}{dt}) = v$.
Hint: Existence of solutions of differential equations.

Exercise 8.9.
Suppose that ∇ and $\tilde{\nabla}$ are two connections on M. For $X, Y \in \Gamma(TM)$, define

$$B(X, Y) = \nabla_X Y - \tilde{\nabla}_X Y.$$

(a) Show that B is tensorial (i.e., it depends only upon $X(p)$ and $Y(p)$).
(b) $B(X, X) = 0$ if and only if ∇ and $\tilde{\nabla}$ have the same geodesics.

Exercise 8.10.
Let (M, g) be a Riemannian manifold, and ∇ is Levi-Civita connection. For a smooth function $f : M \to \mathbb{R}$, define the "Hessian" $\nabla^2 f$ by

$$\nabla^2 f(X, Y) = X(Y(f)) - (\nabla_X Y)(f),$$

where $X, Y \in \Gamma(TM)$. Prove:
(a) $\nabla^2 f$ is symmetric in X and Y.
(b) $\nabla^2 f$ is tensorial.
(c) $\nabla^2 f$ is positive definite if and only if $(f \circ \gamma)'' \geq 0$ for every geodesic γ of ∇.

9 Some Hamiltonian Mechanics

9.1 Calculus of Variations on Manifolds

We now deal with variational principles, which are a basis for the Hamiltonian formulation of mechanics. The material covered here is a generalization of the calculus of variations, which one usually sees in a mathematical physics course. Indeed, instead of dealing with variations in $\mathbb{R}^n$, we study them on manifolds.

In this section, suppose that (M, g) is a Riemannian manifold, and let ∇ denote the corresponding Levi-Civita connection.

Definition 9.1.
A **Lagrangian** L on a manifold M is a C^∞ map $L : TM \to \mathbb{R}$. Given such a function, we can associate a functional

$$\mathcal{F}_{\mathcal{L}} : \{\text{curves } \gamma : [a, b] \to M\} \to \mathbb{R}$$

by

$$\mathcal{F}_{\mathcal{L}}(\gamma) = \int_a^b \frac{1}{2} L(\gamma(t), \dot{\gamma(t)}) dt.$$

One particular functional (and associated Lagrangian) will be of particular importance to us.

Example 9.2.
Define $L(x, v) = \frac{1}{2} g_x(v, v)$. The corresponding functional is given by

$$E(\gamma) = \int_a^b \frac{1}{2} g_{\gamma(t)}(\dot{\gamma(t)}, \dot{\gamma(t)}) dt.$$

This is sometimes known as the **action functional** . In physics, $\frac{1}{2} g_x(v, v)$ often has the interpretation as kinetic energy.

Since the functional associated to a Lagrangian is a function from curves on M into $\mathbb{R}$, it makes sense to ask whether there are curves that maximize or minimize the values of the functional. That is, we want to know whether "critical points" of a functional $\mathcal{F}_{\mathcal{L}}$. Such optimizing curves are said to be L**-critical.**

Definition 9.3.
Given a curve $\gamma : [a, b] \to M$, a smooth variation of γ (with fixed endpoints)

is a smooth map $\Gamma : [a,b] \times (-\epsilon, \epsilon) \to M$ such that
1) $\Gamma(t,0) = \gamma(t)$;
2) $\Gamma(a,s) = \gamma(a)$, $\Gamma(b,s) = \gamma(b)$ for all s.
A curve $\gamma : [a,b] \to M$ is said to be L-critical for a given Langrangian L if for any variation $\gamma_s = \Gamma(t,s)$, we have

$$\frac{d}{ds}\Big|_{s=0} \mathcal{F}_{\mathcal{L}}(\gamma_s) = 0.$$

Remark 9.4.
If $\gamma : [a,b] \to M$ is L-critical, then its restriction $\gamma\Big|_{[a',b']}$ is L-critical for any $[a',b'] \subset [a,b]$.

Our goal is to find necessary conditions for a curve γ to be L-critical for a given Lagrangian L on a manifold M. In looking for such curves, let us assume that γ is such a curve and that $\gamma([a,b])$ lies in a coordinate chart $(U, x^1, \ldots, x^n)$. We also may assume that $U \subset \mathbb{R}^n$, and $x^1, \ldots, x^n, v^1, \ldots, v^n$ are coordinates on $U \times \mathbb{R}^n = TU$. Further, let γ_s be a variation of γ; we have $\gamma_s = (\gamma_s^1, \ldots, \gamma_s^n)$ and $\dot{\gamma_s} = (\gamma_s^1, \ldots, \gamma_s^n, \dot{\gamma_s^1}, \ldots, \dot{\gamma_s^n})$. Let $h(t) = \frac{\partial}{\partial s}\Big|_{s=0} \gamma_s(t)$. Note that $h(a) = h(b) = 0$; otherwise, $h(t)$ is arbitrary. Given $h : [a,b] \to \mathbb{R}^n$, we can take $\gamma_s(t) = \gamma(t) + sh(t)$. Also

$$\begin{aligned} \frac{\partial}{\partial s}\Big|_{s=0} \dot{\gamma_s(t)} &= \frac{\partial}{\partial s}\Big|_{s=0} \Big(\frac{\partial}{\partial t}\Big|_t \gamma_s(t) \Big) \\ &= \frac{\partial}{\partial t}\Big|_t \Big(\frac{\partial}{\partial s}\Big|_{s=0} \gamma_s(t) \Big) \\ &= \frac{\partial}{\partial t}\Big|_t h(t) = \dot{h(t)}. \end{aligned}$$

That is,

$$\frac{\partial}{\partial s}\Big|_t (\gamma_s(t), \dot{\gamma_s(t)}) = (h(t), \dot{h(t)}).$$

Since γ is L-critical,

$$\begin{aligned} 0 &= \frac{d}{ds}\Big|_{s=0} \mathcal{F}_{\mathcal{L}}(\gamma_s) \\ &= \frac{d}{ds}\Big|_{s=0} \int_a^b L(\gamma(s), \dot{\gamma(s)})\, dt \\ &= \int_a^b \frac{\partial}{\partial s}\Big|_{s=0} L(\gamma_s, \dot{\gamma_s})\, dt \end{aligned}$$

$$
\begin{aligned}
&= \int_a^b \Big(\sum_i \frac{\partial L}{\partial x^i}(\gamma_s, \dot{\gamma_s}) \frac{\partial}{\partial s}\Big|_{s=0} \gamma_s^i + \sum_i \frac{\partial L}{\partial v^i}(\gamma_s, \dot{\gamma_s}) \frac{\partial}{\partial s}\Big|_{s=0} \dot{\gamma_s^i} \Big)\ dt \\
&= \int_a^b \Big(\sum_i \frac{\partial L}{\partial x^i} h^i + \sum_i \frac{\partial L}{\partial v^i} \dot{h}^i \Big)\ dt \\
&= \sum_i \int_a^b \frac{\partial L}{\partial x^i} h^i\ dt + \sum_i \frac{\partial L}{\partial v^i} h^i \Big|_a^b - \sum_i \int_a^b \frac{d}{dt}\Big(\frac{\partial L}{\partial v^i}\Big) h^i\ dt \\
&= \sum_i \int_a^b \Big(\frac{\partial L}{\partial x^i} - \frac{d}{dt}\frac{\partial L}{\partial v^i} \Big) h^i\ dt
\end{aligned}
$$

for all $h^1, \dots, h^n$. Thus, we have that

$$
\frac{\partial L}{\partial x^i}(\gamma, \dot{\gamma}) = -\frac{d}{dt}\Big(\frac{\partial L}{\partial v^i}(\gamma, \dot{\gamma})\Big),
$$

the **Euler-Lagrange equations**. This result can be summarized in the following theorem:

Theorem 9.5.
Suppose L is a Lagrangian and that γ is an L-critical curve. Then locally, γ must satisfy the **Euler-Lagrange Equations** :

$$
\frac{\partial L}{\partial x_i}(\gamma, \dot{\gamma}) = \frac{-d}{dt}(\frac{\partial L}{\partial v_i}).
$$

Let us look at a specific case, namely, where L is the energy Lagrangian.

Example 9.6.
Let g be a Riemannian metric and L be the energy Lagrangian on a manifold M; that is, for $x \in T_x M$,

$$
L(x, v) = \frac{1}{2} g_x(v, v).
$$

We can rewrite this expression in coordinates as

$$
L(x, v) = \sum_{k,l} g_{kl}(x) v^k v^l,
$$

whence

$$
\frac{\partial L}{\partial x^i} = \sum_{k,l} \frac{\partial g_{kl}(x)}{\partial x^i} v^k v^l
$$

and

$$\frac{\partial L}{\partial v^i} = \sum_{k,l} (g_{il}(x) v^l + g_{ki} v^k.$$

The Euler-Lagrange equations in this case are

$$\sum_{k,l} \frac{\partial g_{kl}}{\partial x^i} \dot{\gamma}^k \dot{\gamma}^l = -\sum_{k,l} \frac{d}{dt} (g_{il} \dot{v}^l + g_{ki} \dot{\gamma}^k),$$

which implies that

$$\sum_{i,q} g_{iq} \ddot{\gamma}^q = -\frac{1}{2} \sum_{i,k,l} \Big(\frac{\partial g_{ki}}{\partial x^l} + \frac{\partial g_{il}}{\partial x^k} - \frac{\partial g_{kl}}{\partial x^i} \Big) \dot{\gamma}^l \dot{\gamma}^k$$

Letting $(g^{\alpha\beta}) = (g_{\alpha\beta})^{-1}$, we see that $\sum_\beta g^{\alpha\beta} g_{\beta\gamma} = \delta_{\alpha\gamma}$, which implies that

$$\ddot{\gamma}^j = -\sum_{k,l} \Gamma^j_{kl} \dot{\gamma}^k \dot{\gamma}^l,$$

where Γ^i_{kl} are the Christoffel symbols for the Levi-Civita connection. We now see that this is the geodesic equation; thus, L-critical curves for the energy Lagrangian must be geodesics. Let us summarize this as a corollary.

Example 9.7.
Suppose that

$$L(x, v) = \frac{1}{2} g_x(v, v) - V(x),$$

where $V \in C^\infty(M)$ is the potential energy of some mechanical system, and $\frac{1}{2} g_x(v, v)$ represents its kinetic energy. Hamilton's Principle in mechanics states that a particle subject to these forces (in a conservative system) will move exhibit trajectories γ such that the action functional

$$\int_a^b L(\gamma(t), \dot{\gamma}(t))\, dt$$

is minimized. That is, one can obtain physical trajectories of the system by solving the Euler-Lagrange equations with respect to L. Hamilton's Principle is actually equivalent to Newton's Second Law.

Corollary 9.7.1.
If (M, g) is a Riemannian manifold, then curves that are L-critical for the energy Lagrangian are the geodesics for the Levi-Civita connection on M.

The Euler-Lagrange Equations are a local characterization of L-critical curves; we also have another formulation that includes a global statement :

Theorem 9.8.
Suppose that $L : TM \to \mathbb{R}$ is a Lagrangian. Then there is a unique function $E_L : TM \to \mathbb{R}$ and a unique 1-form $\alpha_L \in \Gamma^1(TM)$ which have coordinate expressions

$$E_L = \sum_i v_i \frac{\partial L}{\partial v_i} - L$$

and

$$\alpha_L = \frac{\partial L}{\partial v_i} dx_i.$$

Moreover, $\gamma : [a, b] \to M$, $\dot{\gamma} : [a, b] \to TM$ satisfy the Euler-Lagrange equations if and only if

$$\iota(\ddot{\gamma(t)})(d\alpha_L)_{\dot{\gamma(t)}} = -(dE_L)_{\dot{\gamma(t)}}.$$

The proof of this theorem is non-trivial and proceeds in several steps.
Step 1: Let V be a finite-dimensional vector space. Define a function $R : V \to TV = V \times V$ by $R(v) = (v, v)$. Note that this is a vector field; it is actually the radial vector field, for it has coordinate expression

$$R(v) = \sum_i v^i \frac{\partial}{\partial v^i}.$$

The flow of this vector field is

$$\Phi_t(v) = v \cdot e^t.$$

Now, more generally, if $\pi : E \to M$ is a vector bundle, then

$$\Phi_t(x, v) = (x, e^t v)$$

is a flow, where $x \in M$ and $v \in T_x M$. The corresponding vector field is the radial vector field, which we'll denote simply by R. If $E = TM$ and $(x^1, \dots, x^n)$ are coordinates on M, $(x^1, \dots, x^n, v^1, \dots, v^n)$ are coordinates on TM, and

$$R = \sum_i v^i \frac{\partial}{\partial v^i}.$$

Define

$$E_L = R(L) - L = dL(R) - L.$$

Step 2: The second step is a discussion of the Legendre transform. First, let V be a (finite-dimensional) vector space and $f \in C^\infty(V)$. Then for $v \in V$, $df_v \in V^* \simeq T_v^* V$. That is, there is a map $\mathcal{L}_f : V \to V^*$ given by

$$\mathcal{L}_f(v) = df_v.$$

In coordinates, this is

$$\mathcal{L}_f(v^1, \ldots, v^n) = \Big(\frac{\partial f}{\partial v^1}, \ldots, \frac{\partial f}{\partial v^n}\Big).$$

Now let us extend this definition to an arbitrary manifold M. Given $f \in C^\infty(TM)$, for each $x \in M$, we have

$$\mathcal{L}_f\Big|_{T_xM} : T_xM \to T_x^*M$$

given by

$$v \mapsto d(f\Big|_{T_xM})_v.$$

That is, there is a map $\mathcal{L}_f : TM \to T^*M$ given by

$$\mathcal{L}_f(x,v) = d(f\Big|_{T_xM})_v.$$

In coordinates,

$$\mathcal{L}_f(x,v) = \Big(x^1, \ldots, x^n, \frac{\partial f}{\partial v^1}, \ldots, \frac{\partial f}{\partial v^n}\Big).$$

This concludes step 2, but before we move on, let us see an explicit example of the Legendre transform.

Example 9.9.
Suppose that (M,g) is a Riemannian manifold. Let $L(x,v) = \frac{1}{2}g_x(v,v)$. Recalling how we differentiate bilinear forms, we obtain

$$\begin{aligned} d(L\Big|_{T_xM})_v(w) &= \frac{d}{dt}\Big|_{t=0} L(x, v+tw) \\ &= \frac{d}{dt}\Big|_{t=0} \frac{1}{2} g_x(v+tw, v+tw) \\ &= \frac{d}{dt}\Big|_{t=0} \frac{1}{2}(g_x(v,v) + 2t g_x(v,w) + t^2 g_x(w,w)) \\ &= g_x(v,w) \end{aligned}$$

That is, $\mathcal{L}_L(x,v) = g_x(v, \cdot)$.

Exercise 9.1.
Suppose that (M, g) is a Riemannian manifold. Then g induces a diffeomorphism from TM to T^*M by $(x, v) \mapsto g_x(v, \cdot)$.

Step 3: Next, we discuss the tautological 1-form on the cotangent bundle, T^*M. The content of this discussion is contained in the following lemma:

Lemma 9.10.
Let M be a manifold. Then there is a unique 1-form α on T^*M such that for any coordinates $(x^1, \dots, x^n)$ on M and $(x^1, \dots, x^n, p^1, \dots, p^n)$ on T^*M,

$$\alpha = \sum_i p^i \, dx^i.$$

Proof. Let $q \in M$, $\eta \in T_q^*M$, and $w \in T_{(q,\eta)}(T^*M)$. Define

$$\alpha_{(q,\eta)}(w) = \eta(d\pi_{(q,\eta)}(w)),$$

where $\pi : T^*M \to M$ is the bundle projection, whence $d\pi_{(q,\eta)} : T_{(q,\eta)}(T^*M) \to T_qM$. If $(x^1, \dots, x^n)$ are coordinates on M and $(x^1, \dots, x^n, p^1, \dots, p^n)$ are the corresponding coordinates on T^*M, then $\pi(x^1, \dots, x^n, p^1, \dots, p^n) = (x^1, \dots, x^n)$; hence

$$d\pi_{(q,\eta)}\Big(\frac{\partial}{\partial x^i}\Big|_{(q,\eta)}\Big) = \frac{\partial}{\partial x^i}\Big|_q$$

and

$$d\pi_{(q,\eta)}\Big(\frac{\partial}{\partial p^i}\Big|_{(q,\eta)}\Big) = 0.$$

This in turn means that

$$\begin{aligned} \alpha_{(q,\eta)} &= \sum_i \Big(\alpha_{(q,\eta)}\Big(\frac{\partial}{\partial x^i}\Big|_{(q,\eta)}\Big)dx^i + \alpha_{(q,\eta)}\Big(\frac{\partial}{\partial p^i}\Big|_{(q,\eta)}\Big)dp^i\Big) \\ &= \sum_i \iota\Big(\frac{\partial}{\partial x^i}\Big|_q\Big)dx^i \\ &= \sum_i p^i(\eta)dx^i \end{aligned}$$

That is,

$$\alpha = \sum_i p^i \, dx^i.$$

□

Step 4: Now define

$$\alpha_L = (\mathcal{L}_L)^*\alpha.$$

In coordinates,

$$\mathcal{L}_L(x^1,\ldots,x^n,p^1,\ldots,p^n) = \Big(x^1,\ldots,x^n,\frac{\partial L}{\partial v^1},\ldots,\frac{\partial L}{\partial v^n}\Big),$$

which implies that

$$(\mathcal{L}_L)^*\alpha = \sum_i \frac{\partial L}{\partial v^i}\, dx^i.$$

Step 5: It remains to show that γ solves the Euler-Lagrange equations if and only if

$$\iota(\ddot{\gamma})(d\alpha_L)_{\dot{\gamma}} = -(dE_L)_{\dot{\gamma}}. \tag{4}$$

Throughout this section we use Einstein's summation convention. Before we proceed, recall that $\gamma = (\gamma^1,\ldots,\gamma^n)$, $\dot{\gamma} = (\gamma^1,\ldots,\gamma^n,\dot{\gamma}^1,\ldots,\dot{\gamma}^n)$ and $\ddot{\gamma} = (\gamma^1,\ldots,\gamma^n,\dot{\gamma}^1,\ldots,\dot{\gamma}^n,\ddot{\gamma}^1,\ldots,\ddot{\gamma}^n)$. The Euler-Lagrange equations are

$$\frac{\partial L}{\partial x^i}(\gamma,\dot{\gamma}) = -\frac{d}{dt}\Big(\frac{\partial L}{\partial v^i}(\gamma,\dot{\gamma})\Big) = -\frac{\partial^2 L}{\partial x^j v^i}(\gamma,\dot{\gamma})\dot{\gamma}^j - \frac{\partial^2 L}{\partial v^j v^i}(\gamma,\dot{\gamma})\ddot{\gamma}^j.$$

To prove (4), we first express the equation in coordinates. Noting that

$$\alpha_L = \frac{\partial L}{\partial v^i}\, dx^i$$

and

$$\iota(\ddot{\gamma})\, dx^i \wedge dx^j = \dot{\gamma}^j dx^i - \dot{\gamma}^i dx^j,$$

we see that

$$\begin{aligned}
\iota(\ddot{\gamma})(d\alpha_L)_{\dot{\gamma}} &= \iota(\ddot{\gamma})\Big(\frac{\partial^2 L}{\partial x^j\partial v^i}\, dx^j\wedge dv^i + \frac{\partial^2 L}{\partial v^j\partial v^i}\, dv^j \wedge dx^i\Big) && (5)\\
&= \frac{\partial^2 L}{\partial x^j\partial v^i}(\dot{\gamma}^j\, dx^i - \dot{\gamma}^i\, dx^j) + \frac{\partial^2 L}{\partial v^j\partial v^i}(\ddot{\gamma}^j\, dx^i - \dot{\gamma}^i\, dv^j) && (6)\\
&= \Big(\frac{\partial^2 L}{\partial x^j v^i}\dot{\gamma}^j - \frac{\partial^2 L}{\partial x^i v^j}\dot{\gamma}^j + \frac{\partial^2 L}{\partial v^j v^i}\ddot{\gamma}^j\Big)dx^i - \frac{\partial^2 L}{\partial v^j v^i}\dot{\gamma}^i\, dv^i. && (7)
\end{aligned}$$

Furthermore,

$$\begin{aligned}
dE_L &= d(v^j\frac{\partial}{\partial v^j} - L)\\
&= \frac{\partial}{\partial x^i}\Big(v^i\frac{\partial L}{\partial v^j} - L\Big)\, dx^i + \frac{\partial}{\partial v^i}\Big(v^j\frac{\partial L}{\partial v^j} - L\Big)\, dv^i\\
&= \Big(v^j\frac{\partial^2 L}{\partial x^i v^j} - \frac{\partial L}{\partial x^i}\Big)\, dx^i + \Big(\frac{\partial L}{\partial v^i} + v^j\frac{\partial^2 L}{\partial v^i\partial v^j} - \frac{\partial L}{\partial v^j}\Big)\, dv^i.
\end{aligned}$$

Thus, we obtain

$$-(dE_L)_{\dot\gamma} = \Big(- v^j \frac{\partial^2 L}{\partial x^i v^j} + \frac{\partial L}{\partial x^i}\Big)\, dx^i - \dot\gamma^j \frac{\partial^2 L}{\partial v^i \partial v^j}\, dv^i. \tag{8}$$

Now, equations (7) and (8) are equal if and only if all of their coefficients are equal; that is, if and only if, for all i, we have

$$\frac{\partial^2 L}{\partial x^j \partial v^i}\dot\gamma^j + \frac{\partial^2 L}{\partial v^j \partial v^i}\ddot\gamma^j = \frac{\partial L}{\partial x^i},$$

which is true if and only if γ satisfies the Euler-Lagrange equations.

9.2 Some Symplectic Geometry

This section is meant to serve a two-fold purpose: first, to follow up on the last result of the previous section, and second, to provide a superficial and cursory introduction to symplectic geometry. In the following discussion, V is a finite-dimensional vector space.

Definition 9.11.
A skew-symmetric form $\omega : V \times V \to \mathbb{R}$ is **nondegenerate** (**symplectic**) if, for any $v \in V$, $\omega(v, w) = 0 \forall w$ implies $v = 0$.

Note here that $\omega \in \Lambda^2(V^*)$. In addition, skew symmetry tells us that $\omega(v, v) = -\omega(v, v)$ for all $v \in V$, which means that $\omega(v, v) = 0$ for all $v \in V$. The following lemma establishes an equivalent condition for nondegeneracy.

Example 9.12.
$\omega = dx \wedge dy$ on $\mathbb{R}^2$ is nondegenerate. More generally,

$$\omega = \sum dx^i \wedge dy^i$$

is nondegenerate on $\mathbb{R}^{2n}$, where coordinates are given by $(x^1, \dots, x^n, y^1, \dots, y^n)$.

Lemma 9.13.
$\omega \in \Lambda^2(V^*)$ is nondegenerate if and only if $\omega^\# : V \to V^*$ given by $v \mapsto \omega(v, \cdot)$ is an isomorphism.

Proof. Since V and V^* have the same dimension, $\omega^\#$ is an isomorphism if and only if its kernel is trivial (it is a linear map). Now the kernel of $\omega^\#$ is precisely

$$\{v \in V : \omega(v, w) = 0\ \forall w \in V\}.$$

By the definition of nondegeneracy, this kernel is trivial if and only if ω is nondegenerate. $\square$

Remark 9.14.
Note that

$$\omega^{\#}(v) = \iota(v)\omega.$$

Additionally, note that if ω is nondegenerate, then for $l \in V^*$, there is a unique vector v_l such that

$$\omega(v_l, \cdot) = l(\cdot).$$

That is,

$$\omega(v_l, \cdot) = \iota(v_l)\omega.$$

Definition 9.15.
A vector space V together with a nondegenerate skew-symmetric form ω is called a **symplectic vector space**.

Now we translate this concept to manifolds.

Definition 9.16.
Let Q be a manifold. A 2-form $\omega \in \Omega^2(Q)$ is **symplectic** if
1) $\omega_q \in \Lambda^2(T_q^*Q)$ is nondegenerate for all $q \in Q$;
2) $d\omega = 0$.

Definition 9.17.
A **symplectic manifold** is a manifold M together with a symplectic form ω, written (M, ω).

Example 9.18.
Let $\omega = dx \wedge dy$ and $M = \mathbb{R}^2$. Then (M, ω) is a symplectic form. This construction generalizes naturally to $\mathbb{R}^{2n}$ for any n. In the case $n = 1$, $\omega^{\#}$ can be realized as the matrix

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Let (M, ω) be a symplectic manifold. The nondegeneracy of ω means that, given $f \in C^\infty(M)$, we can implicitly define a vector field X_f corresponding to f. This vector field is called the **Hamiltonian vector field X_f of f**.

Definition 9.19.
Let (M, ω) be a symplectic manifold and suppose that $f \in C^\infty(M)$. The **Hamiltonian vector field X_f of f** is the unique vector field satisfying

$$\iota(X_f)\omega = df.$$

Exercise 9.2.
Check that if ω is a nondegenerate form, then X_f as defined above is indeed unique.

Example 9.20.
Let us now interpret the last result from the previous section. We had that

$$\iota(\ddot{\gamma})d\alpha_L = -dE_L \tag{9}$$

if and only if γ solves the Euler-Lagrange equations. Now, note that if $d\alpha_L$ is nondegenerate, then there is a unique vector field X_{E_L} such that

$$\iota(X_{E_L})d\alpha_L = -dE_L.$$

This says that $\ddot{\gamma}$ is the integral curve of some vector field on the tangent bundle if and only if it solves the Euler-Lagrange equations. In the case of symplectic manifolds, (9) has this particular interpretation.

Example 9.21.
Suppose that

$$L(x, v) = \frac{1}{2}g_x(v, v) - V(x),$$

where $V \in C^\infty(M)$. (If V is taken to represent the potential energy of some mechanical system, then L is the Hamiltonian of that system.) Then the Legendre transform $\mathcal{L}_L : TM \to T^*M$ is a diffeomorphism, and

$$d\alpha_L = d(\mathcal{L}_L^*\alpha) = \mathcal{L}_L^* d\alpha.$$

Now, locally, we have

$$d\alpha = d(\sum_i p_i\, dx_i) = \sum_i dp_i \wedge dx_i,$$

which is nondegenerate. Since $\mathcal{L}_L$ is a diffeomorphism, $d\alpha_L$ is also nondegenerate.

Here are some other observations:
1) If $d\alpha_L$ is nondegenerate, then the energy E_L is conserved. That is, if ϕ_t is the flow of X_L, then

$$E_L(\phi_t(z)) = E_L(z)\ \forall z \in TM,\ \forall t$$

if and only if

$$\begin{aligned}
0 &= \left.\frac{d}{dt}\right|_{t=s} E_L(\phi_t(z)) \\
&= (X_L(E_L))(\phi_t(z)) \\
&= (dE_L)(X_L)(\phi_t(z)) \\
&= (-\iota(X_L)d\alpha_L)(X_L)(\phi_t(z)) \\
&= -(d\alpha_L(X_L, X_L))(\phi_t(z)) \\
&= 0
\end{aligned}$$

since $d\alpha_L$ is skew-symmetric.
2) Symmetries gives us conservation laws. Briefly, we sketch what this means and give an example. First, let X be a complete vector field on M with flow Φ_t. We can lift the flow to TM by defining

$$\Psi_t = d\Phi_t.$$

The chain rule and the group property of Φ_t that Ψ_t satisfies the group property; hence, Ψ_t is a flow on TM. Next, let

$$Y = \left.\frac{d}{dt}\right|_{t=0} \Psi_t;$$

then Y is a vector field on TM.

Definition 9.22.
A complete vector field X on M is an **infinitessimal symmetry** of $L : TM \to \mathbb{R}$ if $\Psi_t^* L = L$, where Ψ_t is the lifted flow to TM. Alternatively, it is an infinitessimal symmetry of L if

$$Y(L) = L,$$

where Y is the lifted vector field to TM.

Here is a theorem of Noether, which we state without proof.

Theorem 9.23.
If X is a symmetry of $L : TM \to \mathbb{R}$, Y is its lifted vector field to TM, and $d\alpha_L$ is nondegenerate, then

$$h_x := \alpha_L(Y) \in C^\infty(TM)$$

is conserved; that is, h is constant along integral curves of X_L, where

$$\iota(X_L)d\alpha_L = -E_L.$$

Example 9.24.
This example shows that rotational symmetry in a central force field (in $\mathbb{R}^3$) gives rise to conservation of angular momentum. Consider a particle in such a field of mass m; by Newton's Second Law, if γ is the trajectory of the particle, then

$$m\frac{d^2}{dt^2}\gamma(t) = F(\gamma, \dot{\gamma}),$$

where F represents the force acting on the particle. If F is conservative, then $F = -\nabla V$ for some potential function $V \in C^\infty(\mathbb{R}^3)$. The appropriate Lagrangian for the system, as mentioned in an earlier example, is

$$L(x, v) = \frac{1}{2}\sum_i (v^i)^2 - V(x).$$

The Euler-Lagrange equations for this Lagrangian state

$$\frac{d}{dt}(m\dot{\gamma}^i) = m\ddot{\gamma}^i = -\frac{\partial V}{\partial x^i}(\gamma).$$

Let us further assume that $V(x) = W(||x||^2)$ for some $W \in C^\infty(\mathbb{R} \setminus \{0\})$. (An example of such a force is gravity.) If A is an orthogonal matrix, then $\langle Ax, Ax\rangle = \langle x, x\rangle$ for any $x \in \mathbb{R}^3$. This in turn means that

$$L(Ax, Av) = \frac{1}{2}m||Av||^2 - V(||x||) = \frac{1}{2}m||v||^2 - V(||x||).$$

Take

$$A(t) = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Now,$\psi_t(x) = A(t)x$ is a flow on $\mathbb{R}^3$. Its lift is

$$\psi_t(x, v) = (A(t)x, A(t)v),$$

which implies that

$$\begin{aligned} \frac{d}{dt}\Big|_{t=0}(A(t))x &= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \begin{pmatrix} x_2 \\ -x_2 \\ 0 \end{pmatrix} = x_2\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_2} \end{aligned}$$

is a symmetry of L. The lifted vector field Y in this case is given by

$$Y = x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2} + v_2 \frac{\partial}{\partial v_1} - v_1 \frac{\partial}{\partial v_2}.$$

Furthermore,

$$\alpha_L = \sum_i \frac{\partial L}{\partial v_i}\, dx_i = \sum_i m v_i\, dx_i.$$

Now, Noether's Theorem tells us that

$$h_x(x, v) = \iota(Y)\alpha_L = m v_1 x_2 - m v_2 x_1$$

is a conserved quantity. This expression, however, is

$$m(v \times x) \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

the x_3 component of angular momentum. Similarly, using other rotations about the x_1 and x_2 axes, we can observe that the other components of angular momentum are conserved, whence the quantity

$$(mv) \times x$$

is conserved.

By similar methods, one can indeed see that translational symmetry yields conservation of linear momentum, as one would expect.

As we mentioned earlier, this has been a very superficial introduction to symplectic geometry and Hamiltonian mechanics. For further study in this area, one needs to concentrate on both Lie groups (which is where the symmetries originate; they often arise as lie groups acting on manifolds) and symplectic geometry.

Appendix A: Multilinear Algebra

In this appendix, we review some of the basic linear algebra concepts necessary for much of the material concerning vector bundles, differential forms, and connections. These concepts include tensor products, exterior algebras, alternating maps, and pairings.

I: Tensor Products of Vector Spaces

Let V and W be finite dimensional vector spaces over $\mathbb{R}$. We want to define a new vector space $V \otimes_{\mathbb{R}} W = V \otimes W$ together with a bilinear map $\otimes : V \times W \to V \otimes W$ such that for any bilinear map $b : V \times W \to Z$, there is a unique linear map $\overline{b} : V \otimes W \to Z$ with $\overline{b}(v \otimes w) = b(v, w)$. This new vector space is called the **tensor product of V and W**, and the existence of $\overline{b}$ is its **universal property**.

Proposition 9.25.
If V and W are two vector spaces and $V \otimes W$ exists, then it is unique (up to isomorphism).

Proof. Suppose that $V \otimes_1 W$ and $V \otimes_2 W$ are two such vector spaces with bilinear maps $\otimes_i : V \times W \to V \otimes_i W$. Then we obtain maps $\overline{\otimes_1}$ and $\overline{\otimes}_2$ such that the following diagram commutes:

$$\begin{tikzcd}
V \otimes_1 W \arrow[dr, shift left, "\overline{\otimes}_2"] & \\
V \times W \arrow[u, "\otimes_1"] \arrow[r, "\otimes_2"'] & V \otimes_2 W \arrow[ul, shift left, "\overline{\otimes}_1"]
\end{tikzcd}$$

Define

$$\begin{aligned}
A_1 &= \overline{\otimes}_1 \circ \overline{\otimes}_2 : V \otimes_1 W \to V \otimes_2 W \\
A_2 &= \overline{\otimes}_2 \circ \overline{\otimes}_1 : V \otimes_2 W \to V \otimes_1 W;
\end{aligned}$$

these are linear maps such that

$$\begin{tikzcd}
V \otimes_i W \arrow[dr, "A_i"] & \\
V \times W \arrow[u, "\otimes_i"] \arrow[r, "\otimes_i"'] & V \otimes_i W
\end{tikzcd}$$

Uniqueness and the fact that the identity maps also make these diagrams commute means that $A_1 = A_2 = id$, whence we get that $V \otimes_1 W \simeq V \otimes_2 W$. $\square$

Proposition 9.26.
Tensor products exist.

Proof. Fix finite dimensional vector spaces V and W (over $\mathbb{R}$). Let $F(V\times W)$ be the vector space spanned by $V \times W$. That is,

$$F(V \times W) = \sum_{(v,w)\in V\times W} \mathbb{R}(v, w).$$

Note that we have an inclusion map $\iota : V \times W \to F(V \times W)$. Now consider the following collection of vectors in $F(V \times W)$:

$$\begin{aligned}
&(v_1 + v_2, w) - (v_1, w) - (v_2, w)\\
&(v, w_1 + w_2) - (v, w_1) - (v, w_2)\\
&\alpha(v, w) - (\alpha v, w)\\
&\alpha(v, w) - (v, \alpha w),
\end{aligned}$$

for all $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$ and $\alpha \in \mathbb{R}$. Let K be the subspace spanned by this collection, and define

$$V \otimes W = F(V \times W)/K.$$

Furthermore, define the map $\otimes : V \times W \to V \otimes W$ by composing

$$V \times W \to F(V \times W) \to F(V \times W)/K.$$

The definition of K forces $\otimes$ to be bilinear; to prove existence, we thus need to verify the universal property.
Suppose $b : V \times W \to Z$ is bilinear. Since $V \times W$ is a basis for $F(V \times W)$, b defines a unique linear map $\overline{b} : F(V \times W) \to Z$ given on the basis by $\overline{b}((v, w)) = b((v, w))$. As b is bilinear, it is 0 on K (by the definition of K); thus, we obtain a unique linear map $\overline{b} : F(V \times W)/K = V \otimes W \to Z$ with $\overline{b}(v \otimes w) = \overline{b}((v, w)) = b((v, w))$. This verifies the universal property. □

If $\{e_1, \dots, e_k\}$ is a basis for V and $\{f_1, \dots, f_j\}$ is a basis for W, then it is evident that $\{e_i \otimes f_j\}$ spans $V \otimes W$. As the next lemma shows, however, more is true: $\{e_i \otimes f_j\}$ is actually a basis for $V \otimes W$.

Lemma 9.27.
If $\{v_i\}$ is a basis for V and $\{w_j\}$ is a basis for W, then $B = \{v_i \otimes w_j\}$ is a basis for $V \otimes W$. In other words, $\dim V \otimes W = \dim W \cdot \dim W$.

Proof. Since we know that B spans $V \otimes W$, we need check only that if

$$\sum_{i,j} c_{ij} v_i \otimes w_j = 0,$$

then $c_{ij} = 0$ for all i and j. Consider $\phi_{kl} : V \times W \to \mathbb{R}$ defined by $\phi_{kl}(v_i, w_j) = 1$ if $(k,l) = (i,j)$ and 0 otherwise. This is a bilinear map, and by the universal property of the tensor product, we obtain linear maps $\overline{\phi_{kl}} : V \otimes W \to \mathbb{R}$. Now,

$$\begin{aligned} 0 &= \overline{\phi_{kl}}\Big(\sum_{i,j} c_{ij} v_i \otimes w_j\Big) \\ &= \sum_{i,j} c_{ij} \cdot \overline{\phi_{kl}}(v_i \otimes w_j) \\ &= \sum_{i,j} c_{ij} \cdot \phi_{kl}(v_i, w_j) \\ &= c_{kl}. \end{aligned}$$

□

Here are some more examples illustrating the power of the universal property.

Example 9.28.
$V \otimes W \simeq W \otimes V$.

Proof. Consider $b : W \times V \to V \otimes W$ defined by

$$b((w,v)) = v \otimes w.$$

Since this map is bilinear, there is a unique linear map $\overline{b} : W \otimes V \to V \otimes W$ with $\overline{b}(w \otimes v) = v \otimes w$. Since $\overline{b}$ is surjective, it is an isomorphism by dimension count. □

Example 9.29.
$V^* \otimes W \simeq \mathrm{Hom}(V, W)$.

Proof. Consider $b : V^* \times W \to \mathrm{Hom}(V, W)$ defined by

$$(b(v^*, w))(v) = v^*(v)w.$$

Since b is bilinear, it induces a linear map $\overline{b} : V^* \otimes W \to \mathrm{Hom}(V, W)$ given by

$$(\overline{b}(v^* \otimes w))(v) = v^*(v)w.$$

Observe that linear maps of the form $b(v^*, w)$ span Hom(V,W) (similarly, rank one matrices span $M_n(\mathbb{R})$). Then $\overline{b}$ is an isomorphism by dimension count. $\square$

Example 9.30.
If $A : V \to W$ and $B : V' \to W'$ are two linear maps, then there is a unique linear map $A \otimes B : V \otimes V' \to W \otimes W'$ such that $(A \otimes B)(v \otimes w) = A(v) \otimes B(w)$.

Proof. Consider $b : V \times W \to V' \otimes W'$ given by

$$b(v, w) = Av \otimes Bw.$$

b is bilinear, whence the universal property guarantees a linear map b : $V \otimes W \to V' \otimes W'$ with the desired property. $\square$

Exercise 9.3.
Show that $V^* \otimes W^* \simeq (V \otimes W)^*$.

Example 9.31.
$V^* \otimes W^* \simeq \text{bilin}(V \times W, \mathbb{R})$.

Proof. Consider $b : V^* \times W^* \to \text{bilin}(V \times W, \mathbb{R})$ given by

$$(b(v^*, w^*))(v, w) = v^*(v) \cdot w^*(w).$$

b is bilinear and hence induces $\overline{b} : V^* \otimes W^* \to \text{bilin}(V \times W, \mathbb{R})$. If $\{v_i^*\}$ is a basis for V^* and $\{w_j^*\}$ is a basis for W^*, then $\{v_i^* \otimes w_j^*\}$ is a basis for $V^* \otimes W^*$ and

$$b(v_i^* \otimes w_j^*) = \phi_{ij},$$

the maps defined above when we proved that $\{v_i^* \otimes w_j^*\}$ is a basis. As the maps $\{\phi_{ij}\}$ form basis for $\text{bilin}(V \times W, \mathbb{R})$, $\overline{b}$ is an isomorphism. $\square$

Exercise 9.4.
$V \otimes (U \otimes W) \simeq (V \otimes U) \otimes W$.

Exercise 9.5.
Prove that given a multilinear map

$$f : \prod_{i=1}^{n} V \to U,$$

then there exists a unique linear map defined on the n-fold tensor product

$$\overline{f} : V^{\otimes n} \to U$$

with

$$\overline{f}(v_1 \otimes \cdots v_n) = f(v_1, \ldots, v_n).$$

II: The Grassman Algebra and Alternating Maps

Definition 9.32.
An **algebra** over $\mathbb{R}$ is a vector space together with a bilinear map $A \times A \to \mathbb{R}$ ("multiplication"). An algebra A is said to be an **algebra with unity** if there is an element $1 \in A$ such that $1 \cdot a = a$ for all $a \in A$.

Example 9.33.
Let M be a manifold. Then $\Gamma(TM)$, the collection of all vector fields on M, is an algebra over $\mathbb{R}$, where the multiplication is given by the Lie bracket. $\Gamma(TM)$ is an example of a non-associative algebra.

Example 9.34.
Fix a finite-dimensional vector space over $\mathbb{R}$. Define $\tau^0(V) = \mathbb{R}$, $\tau^1(V) = V$, and $\tau^k(V) = V^{\otimes k}$, the k-fold tensor product. Next, define

$$\tau(V) = \sum_{k=0}^{\infty} \tau^k(V)$$

(direct sum). We say that $\alpha \in \tau(V)$ has degree k if $\alpha \in \tau^k(V)$. A typical element in $\tau(V)$, however, has the form

$$u = \alpha_1 \oplus \cdots \oplus \alpha_k,$$

where $\alpha_j \in \tau^j(V)$ for $1 \leq j \leq k$. We define a multiplication by noting that if $\alpha \in \tau^k(V)$ and $\beta \in \tau^l(V)$, then $\alpha \otimes \beta \in \tau^{k+1}(V)$.

Remark 9.35.
$\tau(V)$ is an example of a **graded algebra**. An algebra A is said to be **graded by integers** if

$$A = \sum_{i \in \mathbb{Z}} A_i$$

and for $a \in A_i$ and $b \in A_j$, $a \cdot b \in A_{i+j}$.

Definition 9.36. Grassman Algebra
Let V be a finite dimensional vector space over $\mathbb{R}$. The **Grassman algebra** $\Lambda(V)$ is an algebra over $\mathbb{R}$ with unity together with an injective linear map $i : V \to \Lambda(V)$ called the **structure map** which satisfies the following universal property: If A is an algebra over $\mathbb{R}$ with unity and $j : V \to A$ is

a linear map such that $j(v) \cdot j(v) = 0$ for all $v \in V$, then there is a unique algebra map $\overline{j} : \Lambda(V) \to A$ such that the following diagram commutes.

$$\begin{tikzcd} V \arrow[d, "j"'] \arrow[dr, "i"] & \\ A & \Lambda(V) \arrow[l, "\overline{j}"] \end{tikzcd}$$

We now prove that $\Lambda(V)$ actually exists and is unique. Since the proof of uniqueness is very similar to the proof of uniqueness of the tensor product, we concentrate solely on existence.

Proposition 9.37.
If $\Lambda(V)$ exists, then it is unique (up to isomorphism).

Proposition 9.38.
$\Lambda(V)$ exists.

Proof. Let I be the two-sided ideal in $\tau(V)$ generated by the set $\{v \otimes v : v \in V\}$, and define $\Lambda(V) = \tau(V)/I$. Note that since $\tau(V)$ is graded, so is I, and

$$I = \sum_{k=2}^{\infty} I \cap \tau^k(V).$$

Since I is an ideal, $\Lambda(V)$ is an algebra, and the induced multiplication there is denoted by $\wedge$ ("wedge"). So the composition $V \to \Lambda(V) \to \Lambda(V)/I$ is an injection.
Now that we have "constructed" $\Lambda(V)$, let us prove the universal property. Suppose that A is an algebra and that we are given $j : V \to A$ such that $j(v) \cdot j(v) = 0$ for all $v \in V$. Consider a map $b : V \times V \to A$ given by $(v, w) \mapsto j(v) \cdot j(w)$. Since this map is bilinear, there is a unique linear map $j^{(2)} : V \otimes V \to A$ with $j^{(2)}(v \otimes w) = j(v) \cdot j(w)$. By induction, we have $j^{(k)} : V^{\otimes k} \to A$ with

$$j^{(k)}(v_1 \otimes \cdots \otimes v_k) = j(v_1) \cdots j(v_k).$$

In addition, we define $j^{(0)}(a) = a \cdot 1_A$, for all $a \in \mathbb{R}$. In this way, we obtain an algebra map $\tilde{j} : \tau(V) \to A$. By assumption, $\tilde{j}(v \otimes v) = 0$ for all $v \in V$, whence $\tilde{j}\Big|_I = 0$ which implies that there is a map $\overline{j} : \Lambda(V) \to A$ such that $\overline{j}(v) = j(v)$ for all $v \in V$. Since $\overline{j}$ is uniquely determined on generators, it is unique. $\square$

Remark 9.39.
For any $v \in V$, we have $v \wedge v = 0$. Also,

$$0 = (v_1 + v_2) \wedge (v_1 + v_2) = v_1 \wedge v_1 + v_1 \wedge v_2 + v_2 \wedge v_1 + v_2 \wedge v_2$$

gives that

$$v_1 \wedge v_2 = -v_2 \wedge v_1;$$

that is, the wedge product is skew-commutative.

Remark 9.40.
Let $\Lambda^k(V) = \tau^k(V)/(\tau^k(V) \cap I)$ (called the k^{th} **exterior power** of V). Then $\Lambda(V) = \sum_{k=0}^{\infty} \Lambda^k(V)$, and in addition, $\Lambda^0(V) = \mathbb{R}$ and $\Lambda^1(V) = V$. Also, if $\alpha \in \Lambda^k(V)$ and $\beta \in \Lambda^l(V)$, then $\alpha \wedge \beta \in \Lambda^{k+1}(V)$. Thus, $\Lambda(V)$ is also a graded algebra.

Remark 9.41.
We know that if $\{v_1, \ldots, v_n\}$ is a basis for V, then $\{v_i \otimes v_j\}$ is a basis for $V \otimes V$. By induction, $\{v_{i_1} \otimes \cdots \otimes v_{i_k}\}$ is a basis for $V^{\otimes k}$. Thus, $\{v_{i_1} \wedge \cdots \wedge \otimes v_{i_k}\}$ at least generates $\Lambda^k(V)$. Since $\wedge$ is skew-commutative, however, we can reduce this generating set to

$$\{\{v_{i_1} \wedge \cdots \wedge \otimes v_{i_k} : i_1 < \cdots < i_k\},$$

which implies that

$$\Lambda^l(V) = 0$$

whenever $l > \dim V$. Later we will see that the previous spanning set is also a basis.

Definition 9.42. Alternating Maps
A multilinear map

$$f : \prod_{i=1}^{n} V \to \mathbb{R}$$

is said to be alternating if

$$f(v_1, \ldots, v_n) = (-1)^{\sigma} f(v_{\sigma(1)}, \ldots, v_{\sigma(n)}),$$

where $\sigma \in S_n$ is a permutation on n letters. In other words, f is alternating is it is multilinear and if when we permute coordinates, we change the value of f by the sign of the permutation.

Example 9.43.
Consider a vector space V and $a, b \in V^*$. Define

$$(a \wedge b)(v_1, v_2) = a(v_1)b(v_2) - a(v_2)b(v_1).$$

Then $a \wedge b$ is alternating.

Example 9.44.
Let $V = \mathbb{R}^n$. Then $\det : \prod_{i=1}^n \mathbb{R}^n \to \mathbb{R}$ is alternating.

There is a strong connection between alternating maps and exterior powers.

Proposition 9.45. Universal Property of $\Lambda^k(V)$
Let U and V be vector spaces. If $f : V \times \cdots \times V \to U$, there is a unique linear map $\overline{f} : \Lambda^k(V) \to U$ with

$$\overline{f}(v_1 \wedge \cdots \wedge v_k) = f(v_1, \ldots, v_k).$$

Proof. By the universal property of $V^{\otimes k}$, there is a unique linear map $\tilde{f} : V^{\otimes k} \to U$ such that $\tilde{f}(v_1 \otimes \cdots \otimes v_k) = f(v_1, \ldots, v_k)$. Sice f is alternating, $f\Big|_{I \cap V^{\otimes k}} = 0$, where I is the ideal defined in the construction of $\Lambda(V)$. This gives us $\overline{f} : \Lambda^k(V) = V^{\otimes k}/(I \cap V^{\otimes k}) \to U$ with the desired property. $\square$

Corollary 9.45.1.
Let $A : V \to W$ be a linear map. Then there is a unique linear map $\Lambda^k(A) : \Lambda^k(V) \to \Lambda^k(W)$ such that

$$(\Lambda^k(A))(v_1 \wedge \cdots \wedge v_k) = (Av_1) \wedge \cdots \wedge (Av_k).$$

Proof. Consider the map $b : V \times \cdots \times V \to \Lambda^k(W)$ given by

$$(v_1, \ldots, v_k) \mapsto (Av_1) \wedge \cdots \wedge (Av_k).$$

Since b is an alternating map, one may apply the previous proposition to obtain the desired result. $\square$

Lemma 9.46.
Let V be an n-dimensional vector space. Then $\Lambda^n(V) \simeq \mathbb{R}$.

Proof. We may assume that $V = \mathbb{R}^n$. Let $e_1, \ldots, e_n$ be the standard basis; we need to show that $e_1 \wedge \cdots \wedge e_n \neq 0$. $\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ gives that $\det(e_1, \ldots, e_n) = 1$, so that $\overline{\det} : \Lambda^n(\mathbb{R}^n) \to \mathbb{R}$ satisfies $\overline{\det}(e_1 \wedge \cdots \wedge e_n) = 1$, whence $e_1 \wedge \cdot e_n \neq 0$. $\square$

Corollary 9.46.1.
If $\{f_1, \ldots, f_n\}$ is a basis for V, then $\{f_{i_1} \wedge \cdots \wedge f_{i_k} : 1 \leq i_1 < \cdots < i_k \leq n\}$ is a basis for $\Lambda^k(V)$.

Proof. From earlier in the section, we know that the above set generates $\Lambda^k(V)$, so we need only check independence. Suppose

$$0 = \sum a_{i_1,\ldots,i_k} f_{i_1} \wedge \cdots \wedge f_{i_k}.$$

Pick one sequence $j_1 < j_2 < \cdots < j_k$. Let $j_{k+1} < \cdots < j_n$ be the remaining indices. Consider

$$\begin{aligned} & (\sum a_{i_1,\ldots,i_k} f_{i_1} \wedge \cdots \wedge f_{j_k}) \wedge f_{j_{k+1}} \wedge \cdots \wedge f_{j_n} \\ = \quad & a_{j_1,\ldots,j_k} f_{j_1} \wedge \cdots f_{j_k} \wedge f_{j_{k+1}} \wedge \cdots \wedge f_{j_n}, \end{aligned}$$

which gives $a_{j_1,\ldots,j_k} = 0$. □

Corollary 9.46.2.

$$\dim \Lambda^k(V) = \begin{pmatrix} \dim V \\ k \end{pmatrix}$$

III: Pairings

Definition 9.47.
A **pairing** is a bilinear map $\langle \cdot, \cdot \rangle : V \times W \to \mathbb{R}$.

Definition 9.48.
A pairing is **nondegenerate** if

$$\begin{aligned} \langle v_0, w \rangle = 0 \ \forall w \in W \Rightarrow v_0 = 0 \\ \langle v, w_0 \rangle = 0 \ \forall v \in V \Rightarrow w_0 = 0. \end{aligned}$$

Example 9.49.
$\langle \cdot, \cdot \rangle : (V^* \otimes W^*) \times (V \otimes W) \to \mathbb{R}$ given by

$$\langle v^* \otimes w^*, v \otimes w \rangle = v^*(v) \cdot w^*(w)$$

is nondegenerate.

Proposition 9.50.
If $b : V \times W \to \mathbb{R}$ is a nondegenerate pairing, then $V \simeq W^*$ and $W \simeq V^*$.

Proof. Consider $b_1^\# : V \to W^*$ given by $(b_1^\#(v))(w) = b(v,w)$. Then $b_1^\#$ is linear, and $\ker b_1^\# = \{v_0 \in V : b_1^\#(v_0) = 0\} = \{v_0 \in V : b(v_0, w) = 0 \ \forall w\} = \{0\}$. Thus, $\dim V \le \dim W^* = \dim W$. Similarly, we have $\dim W \le \dim V^* = \dim V$. As $\dim W = \dim V$, we then see that $b_1^\#$ is an isomorphism. Similarly, $b_2^\# : W \to V^*$ given by $w \mapsto b(\cdot, w)$ is an isomorphism. $\square$

Proposition 9.51.
There is a nondegenerate pairing $\langle \cdot, \cdot \rangle : \Lambda^k(V^*) \times \Lambda^k(V) \to \mathbb{R}$ *with*

$$\langle v_1^* \wedge \cdots \wedge v_k^*, v_1 \wedge \cdots \wedge v_k \rangle = \det \Big(v_i^*(w_j) \Big).$$

Proof. Consider $b : (V^*)^k \times V^k \to \mathbb{R}$ given by

$$(l_1, \ldots, l_k, v_1, \ldots, v_k) \mapsto \det \Big(l_i(v_j) \Big).$$

For a fixed $(l_1, \ldots, l_k) \in (V^*)^k$, b is alternating in the v's. So there is a map $\overline{b} : (V^*)^k \times \Lambda^k(V) \to \mathbb{R}$ with

$$(l_1, \ldots, l_k, v_1 \wedge \cdots \wedge v_k) \mapsto \det \Big(l_i(v_j) \Big).$$

Similarly, for a fixed $v_1 \wedge \cdots \wedge v_k \in \Lambda^k(V)$, $\overline{b}$ is alternating in the l's, which means that there is a map $\tilde{b} : \Lambda^k(V^*) \times \Lambda^k(V) \to \mathbb{R}$ with the desired property. $\square$

Exercise 9.6.
Suppose that $\dim V = n$. Given a linear map $A : V \to W$, we get a map $\Lambda^n(A) : \Lambda^n(V) \to \Lambda^n(W)$, and since $\Lambda^n(V) \simeq \mathbb{R}$ by the previous lemma, $\Lambda^n(A)$ is multiplication by a scalar. Show that this scalar is $\det A$.

Appendix B: Professor Lerman's Words of Wisdom

–"I have this theory that if I do things slower, you might be able to follow me."
–"On the one hand, I'm not assuming that you know any point-set topology, but on the other hand, I'm going over it so fast that you can't possibly learn it."
–"The faster I go, the faster you should learn."
–"If there is any justice in the world, this should be true."
–"There is no justice in the world...you have to be a professor in order to use the justice argument."

Index