A Brief Survey of Manifolds and Vector Bundles

by

William A. Dulaney

Honors Thesis

Appalachian State University

Submitted to the Department of Mathematical Sciences
in partial fulfillment of the requirements for the degree of

Bachelor of Science

May, 2019

Approved by:

William J. Cook, Ph.D., Thesis Director, Honors Director

Jason Osborne, Ph.D., Second Reader

Eric S. Marland, Ph.D., Chair

Abstract

This is a survey of various key components and machinery used in modern differential geometry. We begin with an introduction to multilinear algebra up to basic notions of tensors and exterior algebras. A survey of basic topology and manifold theory is then presented. This includes a formulation of tangent spaces, vector bundles, and sections on a manifold. This paper then concludes with an introduction to semi-Riemannian geometry including an outline of metrics, connections, and covariant differentiation.

Acknowledgements

I am grateful for the support and guidance given by my advisor Dr. Bill Cook. Without him answering my endless questions and the hours of dedication he put forth for nearly a year, this project would not have been possible.

I would also like to show my gratitude to Dr. Jason Osborne, who has likewise spent hours helping with my intuition on these advanced topics and helped set me in the right direction early on in my explorations of geometry.

Thank you for all of your encouragement.

Contents

1 Introduction **1**

2 Multilinear Algebra **2**

2.1 Vector Spaces and the Dual 2

2.2 Tensors 6

2.3 Exterior Algebra 10

3 Manifolds **11**

3.1 Some Topology 12

3.2 Topological Manifolds 15

3.3 Coordinates 17

3.4 Differentiable Manifolds 17

4 Tangential Structures **24**

4.1 Tangent Spaces on $\mathcal{M}$ 24

4.2 Cotangent Space 28

4.3 Tangent Bundles 30

4.4 Cotangent Bundle 32

4.5 Vector Bundles 33

4.6 Sections 35

5 Semi-Riemannian Geometry **35**

5.1 Lie Algebras and Derivatives 36

5.2 Metrics 37

5.3 Connections 39

5.4 Covariant Derivative 44

1 Introduction

Differential geometry is a cornerstone of mathematics and is a vast and old field of study that was given its first modern manifold based formulation in the 1800's by Bernhard Riemann. Since there has been so much research in this area, getting into the field and playing "catch-up" is quite a challenging endeavor. There is a running joke that once one is able to have a broad understanding of the forefront of differential geometry, they have become too old to apply their knowledge or expand much further. However for many, the field's beauty and the harmony it has with physical reality is seductive and therefore they cannot help but to push forward. Many results over the past century have been influential to not only an exceedingly large span of mathematics, but also to society and humanity's view of the cosmos.

In this paper an attempt is made to give as many viewpoints as possible as we work through each structure that is introduced. Having these multiple vantage points to tackle problems is invaluable for both intuition and progress. Specifically, the bundle approach to differential geometry gives the sense of where one is in the vast space of various structures and mechanisms at play. As for the topics covered, it is assumed the reader has been introduced to linear algebra and has the mathematical maturity of a student who has completed a basic course in analysis and algebra. With this in mind, I have made an attempt to include just enough material for the reader to move through the topics of this paper with the assistance of referenced supplements with the intent to inspire further reading and study.

2 Multilinear Algebra

To begin our venture into geometry we first must review some basic notions of linear spaces. The study of calculus, as the reader should be aware, is a method of analysis via linearization. Therefore, the algebra of linear things, as studied in linear and multilinear algebra, will play a foundational role as we move forward to unfamiliar territory. The reader should be familiar with the notions of linear algebra already, multilinear alegbra is an extension of this to mathematical entities who are linear in "multiple slots." Much of this section on the basics of multilinear algebra is based on material covered in [R] chapters one and two as well as some material references [ON] chapter two and [HK].

2.1 Vector Spaces and the Dual

Before we begin, some possibly unfamiliar notations must be introduced. First, we say the scalar components to a vector are contravariant whereas the basis components are covariant. More on what exactly this means will be discussed later. For now note that contravariant components will be denoted with an upper index as a^i where as covariant components will have a lower index as e_j, for $i, j \in \mathbb{N}$ (the set of positive integers). These indices denote the i-th or j-th piece of some object and do not indicate a power. Beginning in this section, Einstein's summation convention/notation will be introduced as well. In this notation, summation over a given index is implied. This becomes necessary later on as we often have entities that are summed over many indexing sets at once and writing many different sigmas becomes tedious. To see how this works, consider a sum S defined as

$$S = \sum_{i=1}^{n} a^i b_i = a^1 b_1 + a^2 b_2 + ... + a^n b_n \tag{2.1}$$

in the Einstein notation, 2.1 simply becomes

$$S = a^i b_i. \tag{2.2}$$

In order to help the reader become familiar with this, the first few times a mathematical object with a sum is introduced, both forms will be given. From there on, however, the sum will be implied unless otherwise noted. This is a beautiful yet challenging convention to become accustomed to, it generally takes a bit of work for those beginning to study this field to pay attention to context to know what is meant.

Now, recall from linear algebra the definition of a vector space.

Definition 2.1. *A vector space V over a field $\mathbb{F}$ is a set such that, without loss of generality, if $v, w \in V$ (called* vectors*) and $a, b(scalars) \in \mathbb{F}$ then*

1. $av \in V$
2. $a(v + w) = av + aw$
3. $(a + b)v = av + bv$
4. $a(bv) = (ab)v$
5. $1v = v$

Where 1 is the multiplicative identity in $\mathbb{F}$.

As a quick remark, we will only consider finite dimensional vectors spaces over $\mathbb{F} = \mathbb{R}$ (the real numbers) in this discussion, so if the field is not specified assume it is $\mathbb{R}$. However, most of the multilinear algebra we discuss holds over any field.

A subspace of V, denoted $U \subset V$, is a subset where U is also a vector space. If U' is a translation of another subspace of V, say $U' = v + U = \{v + u \mid u \in U\}$, then we say U' is an affine subspace of V. A set of vectors $\{e_i\}$ is said to be linearly independent over $\mathbb{R}$ if for any $a^i \in \mathbb{R}$ we have that

$$\sum_i a^i e_i = 0 \qquad \text{implies} \qquad a_i = 0 \quad \text{for all } i. \tag{2.3}$$

Note that $\sum_i a^i e_i = 0$ can simply be expressed as $a^i e_i = 0$ in the Einstein summation notation. A set of vectors $\mathcal{S}$ is said to span V if every $w \in V$ can be written as a (finite) linear combination of vectors in $\mathcal{S}$. Thus if $\mathcal{S} = \{e_1, \ldots, e_\ell\}$, then the set $\mathcal{S}$ spans

$$V = \text{span}(\mathcal{S}) = \{a^1 e_1 + a^2 e_2 + \cdots + a^\ell e_\ell \mid a^1, a^2, \ldots, a^\ell \in \mathbb{R}\}. \tag{2.4}$$

When β is a linearly independent spanning set for V, we call β a *basis* for V. Given a particular vector space V, all bases have the same cardinality and so if β is a basis for V, then we define $|\beta| = \dim(V)$ (i.e., the *dimension* of V).

For each vector space, there exists what is known as the dual space. First, consider any map

$$f : V \to \mathbb{R}. \tag{2.5}$$

if f is linear, that is

$$f(a^1 v_1 + a^2 v_2) = a^1 f(v_1) + a^2 f(v_2) \quad \text{where } a^i \in \mathbb{R} \text{ and } v_i \in V \tag{2.6}$$

then we call f a linear functional on V.

Definition 2.2. *The set of all linear functionals on V is called the dual space of V. It is denoted as $V^* = \mathrm{Hom}(V, \mathbb{R})$, the set of all (linear) homomorphisms from V into $\mathbb{R}$.*

We can easily conclude that V^* is itself also a vector space. If $s \in \mathbb{R}$ and $f, g \in V^*$ then $sf, f + g \in V^*$ since a scalar times a linear function is also linear and the sum of linear functions is a linear function.

Given a basis $\beta = \{e_i\}$ for V, there exists a canonical dual basis $\beta^* = \{\theta^j\}$ for V^*, also known as a cobasis. The relationship between these two is as follows:

$$\langle e_i, \theta^j \rangle = \langle \theta^j, e_i \rangle = \theta^j(e_i) = \delta_i^j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} . \tag{2.7}$$

The operation $\langle e_i, \theta_j \rangle$ is known as a dual pairing. It is standard convention to not want to distinguish between the e_i's plugged into the θ^j's or vice-versa. This is justified since the double dual V^{**} is naturally isomorphic to V.

Just as with a vector space, elements of the dual space can be described as a linear combination of basis elements. That is. for $f \in V^*$

$$f = \sum_i f_i \theta^i \quad \text{or simply} \quad f = f_i \theta^i \tag{2.8}$$

We will state without proof, though it should be quite obvious, that $\dim(V^*) = \dim(V)$ (for finite dimensional vector spaces). It immediately follows that V is isomorphic to V^* or $V \cong V^*$. In fact if $\dim(V) = n$ then we have $V \cong V^* \cong V^{**} \cong \mathbb{R}^n$ where V^{**} is the double dual of V. It is said that the components of a vector transform contravariantly, since they transform differently than the basis. For a change of basis matrix A^i_j,

$$e'_j = e_i A^i_j \tag{2.9}$$

whereas,

$$(v')^i = (A^{-1})^i_j v^j \tag{2.10}$$

We then observe that the components of a covector transform covariantly, or with the basis elements of V. We see this as follows:

$$(\theta')^i = (A^{-1})^i_j \theta^j \tag{2.11}$$

and

$$f'_j = f_i A^i_j \tag{2.12}$$

Note again that throughout this discussion, unless otherwise suggested, all summed indices are "dummy" indices and are completely arbitrary. Also, repeated indices appearing once as sub- and super-scripts are summed over as per Einstein's convention.

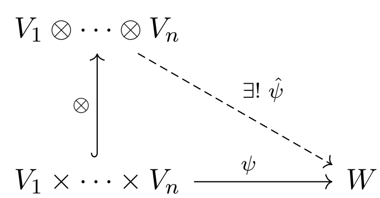
Example 2.3. *Quick covector example: Let* $V = \mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$*. Then* $f(x, y, z) = 5x + 2y - z$ *is a linear map from* $\mathbb{R}^3$ *to* $\mathbb{R}$ *and so* $f \in V^* = (\mathbb{R}^3)^*$*. If we choose the standard basis for* $\mathbb{R}^3$*:* $\beta = \{e_1 = (1, 0, 0), e_2 = (0, 0, 1), e_3 = (0, 0, 1)\}$*, then the dual basis* $\beta^* = \{\theta^1, \theta^2, \theta^3\}$ *consists of the coordinate functions:* $\theta^1(x, y, z) = x$*,* $\theta^2(x, y, z) = y$*, and* $\theta^3(x, y, z) = z$*.*
We find f *can be expressed as a linear combination of these dual basis elements as* $f = 5\theta^1 + 2\theta^2 - \theta^3$*. Now suppose* $v = (1, 0, -1)$*, then* $f(v) = 5(1) + 2(0) - (-1) = 6$*. One can ask: Are we plugging* v *into* f *or* f *into* v*? In some sense, this does not matter, thus the notation:* $\langle f, v\rangle = \langle v, f\rangle = f(v) = 6$*. Again, this reflects the fact that:* $V \cong V^{**}$*.*

2.2 Tensors

Now that linear objects are familiar to us, we may generalize to objects that are linear in more than one ”slot”, that is they are multilinear. Tensors are certain multilinear gadgets. The reader is already familiar with some special cases; specifically a scalar is called a 0-tensor, a vector is a $(1, 0)$ tensor, and a covector is a $(0, 1)$-tensor.

Let $T : V_1 \times \cdots V_n \to W$ be a mapping from several vector spaces $V_1, \ldots, V_n$ to a vector space W. We say T is multilinear (or n-linear) if T is linear in each slot (holding the others constant). In particular, $T(v_1, \ldots, av + w, \ldots, v_n) = aT(v_1, \ldots, v, \ldots, v_n) + T(v_1, \ldots, w, \ldots, v_n)$ for all $v_j \in V_j$ $(i \neq j)$, $v, w \in V_i$, and $a \in \mathbb{R}$.

Tensor products are operations that essentially make multilinear things linear. Consider the following diagram:



Here ψ is multilinear (in n different slots), $\hat{\psi}$ is linear, and the mapping

$$\otimes(v^1, ..., v^n) = v^1 \otimes \cdots \otimes v^n \tag{2.13}$$

is itself multilinear.

Let $v, w \in V$ then we call $v \otimes w$ a 2-tensor, also known as a tensor of order two. Note that the tensor product is not commutative: $v \otimes w \neq w \otimes v$ in general, however it is associative. We say a 0-tensor is a scalar and a 1-tensor is a vector. It is the case however that the set

$$\mathcal{T}^r = \underbrace{V \otimes V \otimes ... \otimes V}_{r-\text{times}} = V^{\otimes r} = \left\{ \sum_{i_1,\ldots,i_r} c_{i_1 \cdots i_r} v_{i_1} \otimes \cdots \otimes v_{i_r} \mid c_{i_1 \cdots i_r} \in \mathbb{R} \text{ and } v_{i_1}, \ldots, v_{i_r} \in V \right\},$$

with $r - \otimes$'s, forms a vector space. It can be shown that the set $\mathcal{T} = \bigcup_{r=0}^{\infty} \mathcal{T}^r$ forms a non-commutative algebra. If T has order r (i.e., $T \in \mathcal{T}^r$) and S has order s, then $\mathcal{T} \otimes \mathcal{S}$ has order $r + s$. We also have for some scalar μ, the following relationships

1. $\mathcal{T} \otimes (\mu \mathcal{S}) = \mu(\mathcal{T} \otimes \mathcal{S})$

2. $\mathcal{R} \otimes (\mathcal{T} + \mathcal{S}) = \mathcal{R} \otimes \mathcal{T} + \mathcal{R} \otimes \mathcal{S}$

3. $(\mathcal{T} + \mathcal{S}) \otimes \mathcal{R} = \mathcal{T} \otimes \mathcal{R} + \mathcal{S} \otimes \mathcal{R}$

Example 2.4. *Let* $\{e_1, e_2, e_3\}$ *be the canonical basis of* $\mathbb{R}^3$*, then the canonical basis for* $\mathbb{R}^3 \otimes \mathbb{R}^3$ *is*

$$\left\{ \begin{array}{ccc} e_1 \otimes e_1, & e_1 \otimes e_2, & e_1 \otimes e_3, \\ e_2 \otimes e_1, & e_2 \otimes e_2, & e_2 \otimes e_3, \\ e_3 \otimes e_1, & e_3 \otimes e_2, & e_3 \otimes e_3 \end{array} \right\}$$

Building from Example 2.4 we have that the most general 2-tensor on $\mathbb{R}^3$ is of the form

$$\mathcal{T} = \sum_{i,j} \mathcal{T}^{ij} e_i \otimes e_j \tag{2.14}$$

Since the basis remains fixed throughout the sum, it is common to refer to $\mathcal{T}$ by its components

$$\mathcal{T} = \mathcal{T}^{ij} \tag{2.15}$$

where the "$e_i \otimes e_j$" are implied. However, this will generally be avoided in this paper as it can be confusing.

Example 2.5. *Let* $v, w \in V$ *then*

$$v \otimes w = \sum_i v^i e_i \otimes \sum_j w^j e_j = \sum_i \sum_j v^i w^j e_i \otimes e_j = v^i w^j e_i \otimes e_j \tag{2.16}$$

As it turns out, the definition of a tensor, as outlined above, is too restrictive and does not account for dual spaces. In all of its simplicity, we define a general tensor as

$$T = T^{i_1 i_2 \ldots i_r}_{j_1 j_2 \ldots j_s} e_{i_1} \otimes e_{i_2} \otimes \ldots \otimes e_{i_r} \otimes \theta^{j_1} \otimes \theta^{j_2} \otimes \ldots \otimes \theta^{j_s} \tag{2.17}$$

where the $T^{i_1 i_2 \ldots i_r}_{j_1 j_2 \ldots j_s}$ are the components and the summation over each sub-index is implied. This multiindexing might seem daunting at first, however we are simply allowing for any possibility of indexing. Imagine working in $\mathbb{R}^3 \otimes \mathbb{R}^3$, then our indexing set could just as well be $\{1, 2, 3\}$ for both i and j. Nonetheless, to avoid a cluster of indicies, we introduce a multiindex notation, such that 2.17 becomes

$$T = T^{\mathbf{I}}_{\mathbf{J}}\, e_{\mathbf{I}} \otimes \theta^{\mathbf{J}}. \tag{2.18}$$

Given a tensor, as depicted in equation 2.17, we say T is a (r, s)-tensor. Furthermore, if T is an (r, s)-tensor and S is a (p, q)-tensor then $T \otimes S$ is a $(r + p, s + q)$-tensor as expected. Tensors can be viewed as multilinear maps, meaning they are linear in each slot. Therefore a (r, s) tensor can be expressed as

$$T : \underbrace{V^* \times V^* \times ... \times V^*}_{r-\text{times}} \times \underbrace{V \times V \times ... \times V}_{s-\text{times}} \to \mathbb{R} \tag{2.19}$$

where for $v, u, w \in V$ or V^* as appropriate and $a, b \in \mathbb{R}$.

$$T(v_1, ..., au + bw, ..., v_{r+s}) = aT(v_1, ..., u, ..., v_{r+s}) + bT(v_1, ..., w, ..., v_{r+s}). \tag{2.20}$$

There are many ways to interpret a given tensor. In practice, we can do what is called currying. To see how this unfolds, consider the following example.

Example 2.6. *For the mapping $T : (V^*)^{\otimes 3} \otimes V \to \mathbb{R}$, we know T is a $(3, 1)$-tensor. It can also be interpreted as mappings:*

$$T : V^* \otimes V \to V \otimes V, \quad T : (V^*)^{\otimes 2} \otimes V \to V, \quad T : (V^*)^{\otimes 2} \to V^* \otimes V,$$

$$T : V \to V^{\otimes 3}, \qquad T : \mathbb{R} \to V^* \otimes V^{\otimes 3}, \quad \textit{and as} \quad T \in V^{\otimes 3} \otimes V^*.$$

For example, considering $T : (V^*)^{\otimes 3} \otimes V \to \mathbb{R}$ and selecting $f \otimes v \in V^* \otimes V$, we have $T(f, \cdot, \cdot, v)$ leaves two dual vector slots unevaluated so that $T(f, \cdot, \cdot, v) : (V^*)^{(\otimes 2)} \to \mathbb{R}$ which corresponds with $V^{\otimes 2}$. Thus the interpretation $T : V^* \otimes V \to V \otimes V$. In components this looks like: $T^{ijk}_{\ell} f_i v^{\ell}$.

For a general tensor T, that is n times contravariant and m times covariant, the contraction over a pair of indicies i and j for $T = T^{\cdots i \cdots}_{\cdots j \cdots}$ yields a tensor that is $n - 1$ times contravariant and $m - 1$ times covariant. Consider for simplicity the $(1,1)$-tensor $T : V^* \times V \to \mathbb{R}$. If we consider the corresponding map $T : V \to V$, then the contraction of T is simply the trace, such that

$$\text{contraction of } (T) = \text{contraction of } (T^i_j e_i \otimes \theta^j) = T^i_i \tag{2.21}$$

We now explore how the indices contract or sum when an input is given for a general tensor and to tie together our discussion.

Example 2.7. *Consider (r,s)-tensors on V. Fix bases for V and V^*: $\beta = \{e_i\}$ and $\beta^* = \{\theta^j\}$ respectively. Then the space of (r,s)-tensors $\mathcal{T}^{\otimes r}_{\otimes s}$ is defined as*

$$\mathcal{T}^{\otimes r}_{\otimes s} = V^{\otimes r} \otimes (V^*)^{\otimes s} \cong \tilde{\mathcal{T}} = \{T : (V^*)^{\otimes r} \otimes V^{\otimes s} \to \mathbb{R} \mid T \textit{ is linear}\}.$$

Then for $v^{j_k}, f_{i_\ell} \in \mathbb{R}$ and contracting over all elements of β and β^*

$$\begin{aligned}
&\left(T^{\mathbf{I}}_{\mathbf{J}}\, e_{\mathbf{I}} \otimes \theta^{\mathbf{J}}\right)(f_{k_1}\theta^{k_1}, ..., f_{k_r}\theta^{k_r}, v^{\ell_1}e_{\ell^1}, ..., v^{\ell_s}e_{\ell_s}) \\
&= f_{k_1}\cdots f_{k_r} v^{\ell_1}\cdots v^{\ell_s} T^{i_1\cdots i_r}_{j_1\cdots j_s} e_{i_1}(\theta^{k_1})\cdots e_{i_r}(\theta^{k_r})\theta^{j_1}(e_{\ell_1})\cdots\theta^{j_s}(e_{\ell_s}) \\
&= f_{k_1}\cdots f_{k_r} v^{\ell_1}\cdots v^{\ell_s} T^{i_1\cdots i_r}_{j_1\cdots j_s} \delta^{k_1}_{i_1}\cdots\delta^{k_r}_{i_r}\delta^{j_1}_{\ell_1}\cdots\delta^{j_s}_{\ell_s} \\
&= T^{i_1\cdots i_r}_{j_1\cdots j_s} f_{i_1}...f_{i_r} v^{j_1}...v^{j_s} = \textit{Constant} \in \mathbb{R}.
\end{aligned}$$

At this point it is fair to note that tensors can have a great deal of symmetry to them. For simplicity of notation we will first consider a $(0,2)$-tensor. If $T_{ij} = T_{ji}$ then we say T is symmetric, otherwise if $T_{ij} = -T_{ji}$ then T is called anti-symmetric. It is easy to show that the symmetric part of T can be expressed as

$$T_{(ij)} = (T_{sym})_{ij} = \frac{1}{2}(T_{ij} + T_{ji}) \tag{2.22}$$

and the anti-symmetric part can be expressed as

$$T_{[ij]} = (T_{asym})_{ij} = \frac{1}{2}(T_{ij} - T_{ji}) \tag{2.23}$$

so that $T = T_{sym} + T_{asym}$.

For the following discussion, call Perm_p the set of all permutations on $\{1, \dots, p\}$. For a permutation $\sigma \in \mathrm{Perm}_p$, denote its sign by $(-1)^\sigma = \mathrm{sign}(\sigma)$ where $(-1)^\sigma = 1$ if σ is even (can be written as an even number of transpositions) and $(-1)^\sigma = -1$ if σ is odd (can be written as an odd number of transpositions). A $(0, p)$-tensor is said to be totally symmetric if

$$T_{i_{\sigma(1)} i_{\sigma(2)} \dots i_{\sigma(p)}} = T_{i_1 i_2 \dots i_p} \tag{2.24}$$

for every $\sigma \in \mathrm{Perm}_p$. On the other hand, we say T is totally anti-symmetric, or skew symmetric or alternating, if

$$T_{i_{\sigma(1)} i_{\sigma(2)} \dots i_{\sigma(p)}} = (-1)^\sigma T_{i_1 i_2 \dots i_p} \tag{2.25}$$

again for all $\sigma \in \mathrm{Perm}_p$. These notions can be extended to $(p, 0)$ and even (r, s) tensors, but this is beyond the scope of this paper. See [R] for more details.

2.3 Exterior Algebra

Exterior algebra or Grassmann algebra is a stand alone theory focusing on alternating tensors. We will soon use this machinery to develop the notion of a differential form. To begin we consider the general space of p-vectors on V, denoted as $\bigwedge^p V$, whose basis is given by all p-vectors of the form $e_I = e_{i_1} \wedge \cdots \wedge e_{i_p}$.

Definition 2.8. *The exterior algebra of V, denoted $\bigwedge V$, is the direct sum of all $\bigwedge^p V$ where $p \in \mathbb{N}$.*

The exterior algebra is a graded algebra. In particular, it is a vector space equipped with a multiplication, $\wedge$, and it is the direct sum of subspaces: $\bigwedge^p V$. Moreover, the following relations are satisfied:

1. $\lambda \wedge (a\mu + b\nu) = a\lambda \wedge \mu + b\lambda \wedge \nu$ for $\lambda \in \bigwedge^k V$, $\mu, \nu \in \bigwedge^k V$, and $a, b \in \mathbb{R}$.

2. $(a\mu + b\nu) \wedge \lambda = a\mu \wedge \lambda + b\nu \wedge \lambda$ for $\lambda \in \bigwedge^k V$, $\mu, \nu \in \bigwedge^k V$, and $a, b \in \mathbb{R}$.

3. $\lambda \wedge (\mu \wedge \nu) = (\lambda \wedge \mu) \wedge \nu$ for $\lambda \in \bigwedge^k V$, $\mu \in \bigwedge^\ell V$, and $\nu \in \bigwedge^m V$.

4. Given $\lambda \in \bigwedge^k V$ and $\mu \in \bigwedge^\ell V$, then $\lambda \wedge \mu \in \bigwedge^{k+\ell} V$

5. $\mu \wedge \lambda = (-1)^{k\ell} \lambda \wedge \mu$, for $\lambda \in \bigwedge^k V$ and $\mu \in \bigwedge^\ell V$ (graded anti-symmetry,)

6. $\lambda \wedge \lambda = 0$ if $\lambda \in \bigwedge^k V$ and k is odd.

It is worth noting the following combinatorial results, given $\dim(V) = n$ then we have the following:

Graded Component	Basis	Dimension
$\bigwedge^0 V = \mathbb{R}$	$\{1\}$	$\binom{n}{0} = 1$
$\bigwedge^1 V = V$	$\{v_1, \ldots, v_n\}$	$\binom{n}{1} = n$
$\bigwedge^2 V$	$\{v_i \wedge v_j \mid i < j\}$	$\binom{n}{2}$
$\ldots$	$\ldots$	$\ldots$
$\bigwedge^i V$	$\{v_{k1} \wedge ... \wedge v_{k_i} \mid k_1 < k_2 < \cdots < k_i\}$	$\binom{n}{i}$
$\ldots$	$\ldots$	$\ldots$
$\bigwedge^n V$	$\{v_1 \wedge v_2 \wedge \cdots \wedge v_n\}$	$\binom{n}{n} = 1$

Example 2.9. *Consider $v, w \in V$ the wedge product can be defined as $v \wedge w = v \otimes w - w \otimes v$, which is a 2-vector that lives in the space $\bigwedge^2 V$. If $v = v^i e_i$ and $w = w^j e_j$ then*

$$v \wedge w = v \otimes w - w \otimes v = v^i e_i \otimes w^j e_j - w^j e_j \otimes v^i e_i$$

$$= v^i w^j (e_i \otimes e_j - e_j \otimes e_i) = v^i w^j e_i \wedge e_j.$$

Using an algebraic approach greatly simplifies computations with p-vectors, which are the anti-symmetric tensors from the end of last section, so much that we can avoid tensor products all together. We will return to $\bigwedge V$ later to discuss differential forms, but first we need the notion of a manifold.

3 Manifolds

As we move to more geometric notions, we must define a structure suitable for calculus. As it turns out, topology gives us the minimum structure required for continuity which allows us to codify the notion of a homeomorphism. On top of this structure, a manifold structure is required to do calculus in a familiar way. Both topology and manifold theory allow us to relate local spaces to a Euclidean structure and create a necessary foundation to pursue differential geometry, here we will pick through the most important aspects of each. This section is based heavily on material from [M] chapters one through seven and [R] chapter three.

3.1 Some Topology

What is topology? In a sense, topology allows one to take a bare set and equip it with the bare minimum amount of mechanisms to discuss and prove results involving continuity. A more general mathematical entity often allows for more powerful theorems and better outlines the underlying structures at hand. For a detailed introduction of topology see [M]. We now present the definition of a topology.

Definition 3.1. *Given an arbitrary set X, a topology on X is a collection $\mathcal{O}$ of subsets of X for which the following properties are satisfied:*

1. *$\emptyset$ and X are in $\mathcal{O}$.*

2. *Any arbitrary union of elements of $\mathcal{O}$ must be in $\mathcal{O}$.*

3. *The intersections of finitely many elements of $\mathcal{O}$ must also be in $\mathcal{O}$.*

Given that a topology has been specified for a set X, we may call X a topological space. A subset $U \subset X$ is said to be open if $U \in \mathcal{O}$. If the complement of a set C, denoted $X - C$, is open then we say C is closed. It would be helpful to note that a set can be open and closed at the same time, as well as neither.

Example 3.2. *Claim: For any set X, the power set, $P(X)$ is a topology on X.*

Proof: If $P(X)$ is the power set of X (i.e., if the set $\mu \subseteq X$ then $\mu \in P(X)$), then by definition X and $\emptyset$ are contained within $P(X)$. Since $P(X)$ is the set of all subsets of X, for any collection of subsets $\alpha_i \subseteq X$ indexed by $i \in I$, it follows in general that $\bigcap_{i \in I} \alpha_i \in P(X)$. Thus not only finite intersection, but also arbitrary intersections lie in $P(X)$. Additionally, we also have that $\bigcup_{i \in I} \alpha_i \in P(X)$ so that arbitrary unions are contained in $P(X)$. Thus $P(X)$ is a topology on X.

Definition 3.3. *The standard topology on $\mathbb{R}$ consists of open sets of the form $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$ as well as arbitrary unions of such sets.*

From this definition, it is trivial to see that

$$\bigcup_{n \in E} (-n, n) = \mathbb{R}$$

for many indexing sets, E, such as and certainly not limited to $\mathbb{N}$ and $\mathbb{R}_{>0}$.

Example 3.4. *Consider $\mathbb{R}$ with the standard topology and the sets $(a, b), (a, b], [a, b]$ for $a, b \in \mathbb{R}$. By definition, the set (a, b) is open in $\mathbb{R}$ but observe that*

$$\mathbb{R} - (a, b) = (-\infty, a] \cup [b, \infty)$$

cannot be written as a union of open intervals so that it is not open in the topology, hence (a, b) is not closed. Similar analysis of the other sets would tell us that $(a, b]$ is neither open nor closed and $[a, b]$ is closed but not open.

Let us now explore the space $\mathbb{R}^n$ in more detail as it will be very important to all future discussions. First, note that $\mathbb{R}^n$ is a vector space (we can add and scale its elements). Let $\mathbf{a} = \langle a^1, \ldots, a^n \rangle, \mathbf{b} = \langle b^1, \ldots, b^n \rangle \in \mathbb{R}^n$ and recall the dot product defined as $\mathbf{a} \bullet \mathbf{b} = a^1 b^1 + a^2 b^2 + \ldots + a^n b^n$. Having equipped $\mathbb{R}^n$ with the dot product, we have given $\mathbb{R}^n$ the structure of an inner product space. Now define the norm on $\mathbb{R}^n$ as $\|\mathbf{v}\| = \sqrt{\mathbf{v} \bullet \mathbf{v}}$. This gives $\mathbb{R}^n$ a normed space structure. For a more detailed description of inner product spaces, normed spaces, and other related material see [K] chapters one through three.

We now have the machinery to define a metric on $\mathbb{R}^n$, known as the standard Euclidean distance, we define this as follows:

$$d(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\| = \sqrt{(v^1 - w^1)^2 + \ldots + (v^n - w^n)^2} \tag{3.1}$$

Think of this as a more general Pythagorean Theorem. Topologically speaking, we have the following precise definition of a metric.

Definition 3.5. *A metric on a set X is a function*

$$d : X \times X \to \mathbb{R}$$

Where the following properties are satisfied:

1. $d(\mathbf{x}, \mathbf{y}) \geq 0$ *for all* $\mathbf{x}, \mathbf{y} \in X$*;* $d(\mathbf{x}, \mathbf{y}) = 0$ *if and only if* $\mathbf{x} = \mathbf{y}$.

2. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ *for all* $\mathbf{x}, \mathbf{y} \in X$.

3. $d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) \geq d(\mathbf{x}, \mathbf{z})$ *for all* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$*, that is we have a triangle inequality.*

If X *is equipped with a metric, we call* X *a* metric space*.*

We note that $\mathbb{R}^n$ equipped with the Euclidean metric gives it the structure of a metric space. A metric for a metric space is a generalization of how to measure distances. When we discuss metrics later in this paper, the later kind of metric, a metric tensor on a manifold, is more-or-less an inner product on tangent spaces. So while there may be similarities, the geometric metrics on a manifold are quite different from the topological kind of metrics on a metric spaces which we defined here.

Just as open sets in $\mathbb{R}$ are built from open intervals, open sets in $\mathbb{R}^n$ are built from open balls.

Definition 3.6. *An open ball, sometimes called an* ϵ*-ball, centered at a point* $\mathbf{x}_0 \in \mathbb{R}^n$ *with radius* $\epsilon > 0$ *is defined as*

$$\mathcal{B}_\epsilon(\mathbf{x}_0) = \left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; \|\mathbf{x} - \mathbf{x}_0\| < \epsilon \right\}.$$

Observe that for a given $x_0 \in \mathbb{R}$ and $\epsilon > 0$, $\mathcal{B}_\epsilon(x_0)$ generates an open interval containing x_0, namely $(x_0 - \epsilon, x_0 + \epsilon)$. Now define an open set in $\mathbb{R}^n$ to be a union of open balls. Equivalently a set U is open in $\mathbb{R}^n$ if given some $\mathbf{x} \in U$ there is some $\epsilon > 0$ such that $B_\epsilon(\mathbf{x})$ is contained in U. Now define an open neighborhood of $\mathbf{x}_0$ to be an open set U with $\mathbf{x}_0 \in U$. For example, $B_\epsilon(\mathbf{x})$ is an open neighborhood of $\mathbf{x}$.

With the concept of open neighborhoods in play, we can begin to discuss what it would require topologically for a function to be continuous.

Definition 3.7. *Let a function* $f : X \to Y$*, where* X *and* Y *are topological spaces and let* $U \subseteq X$ *and* $V \subseteq Y$ *be open sets in* X *and* Y *respectively. We say* f *is continuous if for each open* V *in* Y*,* $f^{-1}(V) = \{x \in X \mid f(x) \in V\}$ *is open in* X*.*

One can define continuity of f at a point x_0 as follows; for every open neighborhood V of $f(x_0)$ (i.e., V is an open set with $f(x_0)$ in V) there exists some open neighborhood of x_0, say U, in X such that $f(U)$ is contained in V. It is straightforward to prove that f is continuous on X if and only if f is continuous at every point in X.

This definition of continuity is surprisingly simple and may seem suspicious at first. However, though it will be omitted here, it can be proven that this definition is equivalent to the standard ϵ-δ-definition used in analysis. Building on this topological notion of continuity, we will require a function to have stronger property to be sufficient for differentiability. An equivalence of differentiable structures is called a diffeomorphism which will be discussed later. Here we define the notion of an equivalence of continuous (or topological) structures which is called a homeomorphism. Homeomorphisms preserve all topological properties.

Definition 3.8. *Let X and Y be topological spaces and consider a bijective function $f : X \to Y$. In the case that both $f : X \to Y$ and $f^{-1} : Y \to X$ are continuous, f is called a homeomorphism and we say that X and Y are homeomorphic topological spaces.*

To have a useful geometry for calculus, the ability to separate two points by non overlapping neighborhoods is also required. In topology, there exists a whole host of what are known as separation or T_i axioms. For our purposes a Hausdorff or T_2 space will be sufficient.

Definition 3.9. *A Hausdorff space is defined as a topological space X in which for each distinct pair of points x_1, x_2 there exists open neighborhoods $x_1 \in U_1$ and $x_2 \in U_2$ such that $U_1 \cap U_2 = \emptyset$.*

It can be easily proven that $\mathbb{R}^n$ is a Hausdorff space. As a sketch of this proof, note that any two points $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$, can be separated by ϵ-ball's whose radius is less than $\frac{1}{2}\|\mathbf{x}_2 - \mathbf{x}_1\|$ and are centered at $\mathbf{x}_1$ and $\mathbf{x}_2$ respectively.

3.2 Topological Manifolds

The next natural step in our discussion, as we move toward higher dimensions, is to introduce the idea of a topological manifold. Topological manifolds are familiar structures similar to that of a curve or a surface. To start, we say a space X is locally Euclidean if there exists an $n \geq 0$ such that at each point in X there is an open neighborhood that is homeomorphic to some open subset $U \subset \mathbb{R}^n$.

Definition 3.10. *A topological manifold is a set $\mathcal{E}$ that is a locally Euclidean Hausdorff space.*

Unraveling this definition, we have $(\mathcal{E}, \mathcal{O})$ is a topological n-dimensional manifold if $\mathcal{E}$ is a Hausdorff space and for all points p in $\mathcal{E}$ there exists an open set $\mathcal{U}$ in the topology $\mathcal{O}$ with $p \in \mathcal{U}$ and a mapping $g : \mathcal{U} \to g(\mathcal{U}) \subset \mathbb{R}^n$ for which g must satisfy:

1. $g(\mathcal{U})$ is open in $\mathbb{R}^n$
2. g is invertible
3. g is continuous
4. g^{-1} is continuous

That is to say, g is a homeomorphism between an open neighborhood of p and an open set in $\mathbb{R}^n$. To build intuition of what a manifold is, consider the following examples:

Example 3.11. *A cone is an example of a 2-dimensional topological manifold. We can use a single map g for all points. Think of a plane tangent to the vertex of the cone and let g be the map which projects the cone down onto the plane. This map will give a homeomorphism between the cone and $\mathbb{R}^2$.*

On the other hand, a figure-eight is not a topological manifold; since there exists an intersection in the center. If we try to take any open neighborhood of the intersection, we get a cross and there is no way to map that in a homeomorphic way onto an open subset of $\mathbb{R}$.

Sketch of proof that figure-eight is not a manifold: Take a part of a loop. Clearly this is homeomorphic to an interval in $\mathbb{R}^1$. So if this is to be a manifold, it must be a 1-dimensional manifold.

Let U be a neighborhood of the cross say at x_0. Let $g : U \to V$ where V is some open subset of $\mathbb{R}$. Then g restricts to a homeomorphism from $U - \{x_0\}$ to $V - \{g(x_0)\}$. But $U - \{x_0\}$ (the cross with intersection removed) has 4 connected components while $V - \{g(x_0)\}$ (an interval with a point removed) has 2 connected components. But continuous maps send connected sets to connected sets and *thus must preserve the number of path components (contradiction). Therefore, there is no such g and so the figure-eight is not a manifold.*

3.3 Coordinates

Before we move forward, we must define how we view coordinates on open sets. Given a point $x \in U \subset \mathbb{R}^n$ (an open neighborhood) the coordinates of x are given as $(x^1, ..., x^n)$, where each x^i is called a component of coordinates of the point. At the same time, we also often view the coordinates as coordinate functions instead, defined as

$$x^i : \mathbb{R}^n \to \mathbb{R} \tag{3.2}$$

where for some point $a = (a_1, ..., a_n) \to a_i$ Equivalently we have that $x^i(a) = a_i$. This should right away seem suspicious as each x^i can be viewed as either a function or the i-th component of a point. Nonetheless this is a standard abuse of notation and we will rush forward with it. Note also that $\frac{\partial x^j}{\partial x^i} = \delta^j_i$.

Note that a bjiective function $g : U \to V$ between two open subsets of $\mathbb{R}^n$, say $g(x^1, \ldots, x^n) = (y^1(x^1, \ldots, x^n), \ldots, y^n(x^1, \ldots, x^n))$ is called a diffeomorphism if g is a homeomorphism and each $y^i(x^1, \ldots, x^n)$ has partial derivatives of all orders defined on U.

Definition 3.12. *If a mapping $f : U \to V$ is a homeomorphism and both f and f^{-1} are smooth, then we say f is a diffeomorphism.*

3.4 Differentiable Manifolds

Again, in order to arrive at a structure necessary for calculus, we need to have a space that locally looks familiar to Euclidean space and is compatible when transitioning between these neighborhoods. To be precise, in order to achieve this we take the following definition of a manifold:

Definition 3.13. *An n-dimensional smooth manifold $\mathcal{M}$ is a Hausdorff topological space with a countable collection of opens sets $\{U_i\}$, called coordinate neighborhoods or coordinate patches, that cover $\mathcal{M}$ and a collection of maps ϕ_i, called coordinate maps (or charts), that satisfy:*

1. *Each $\phi_i : U \to \mathbb{R}^n$ is a homeomorphism onto an open subset of $\mathbb{R}^n$ ($\mathcal{M}$ is locally Euclidean.)*

2. *If U_i and U_j are two overlapping coordinate neighborhoods with coordinate maps ϕ_i and ϕ_j, then $\phi_j \circ \phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$ is a diffeomorphism in $\mathbb{R}^n$. (Coordinates are compatible on the overlaps.)*

Each pair (U_i, ϕ_i) is called a coordinate chart, the collection of coordinate charts form what is called an atlas and $\phi_j \circ \phi_i^{-1}$ are called the transition functions of the atlas. This can be visualized in the figure below.

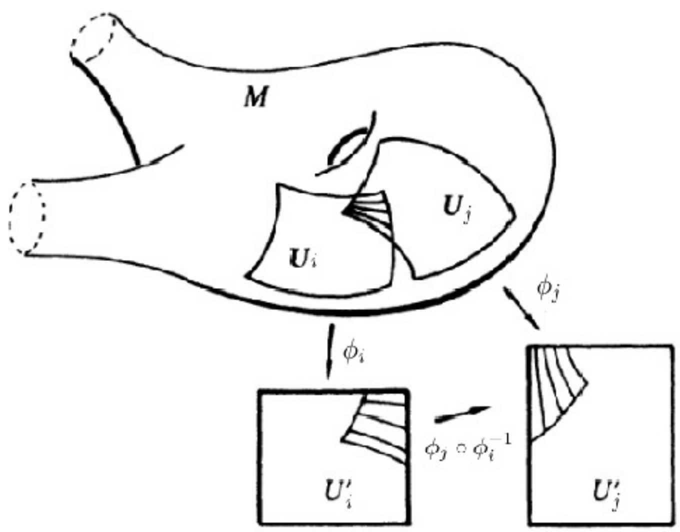


Figure 1: Depiction of coordinate charts and transitions on a manifold. [NLab]

It is also worth noting that every atlas is contained within a maximal atlas. Let $\mathfrak{A}_1$ and $\mathfrak{A}_2$ be maximal atlases, then

$$\mathfrak{A}_1 = \mathfrak{A}_2 \qquad \text{or} \qquad \mathfrak{A}_1 \cap \mathfrak{A}_2 = \emptyset. \tag{3.3}$$

To see how these mechanisms unfold, let (U, ϕ) be a coordinate chart and suppose $\phi(p) = q$. If $x^1, ..., x^n$ are the standard coordinates on $\mathbb{R}^n$ then we have for the point p, its ϕ-coordinates are $\phi(p) = (x^1(q), ..., x^n(q))$. Again, each x^i is called a component of the coordinates of the point. Recall that coordinates can be viewed as functions: $x^i : U \to \mathbb{R}$. So for $p \in \mathcal{M}$, if $\phi(p) = (p_1, ..., p_n) \in \phi(U) \subseteq \mathbb{R}^n$, then $x^i(p) = p_i$.

If $\mathcal{M}$ is a smooth manifold, that is the manifold is globally infinitely differentiable, we call it a C^∞ manifold. As a side note, given a function or a manifold, we say it is C^m if it or its transition functions are m-times continuously differentiable. So in particular, C^0 just requires continuity and so a C^0 manifold is a topological manifold. It follows that $C^0 \subsetneq C^1 \subsetneq ...C^\infty \subsetneq C^\omega$ where C^ω is denotes real analytic functions.

Example 3.14. *A circle is a manifold: Let $S^1 = \{(x, y) \mid x^2 + y^2 = 1\}$. The reader should note that there are many ways in which one can cover a given space with open sets. To cover the circle, we will be using the following:*

- $\phi^+ : \{(x, y) \in S^1 \mid x > 0\} \to \{y \mid -1 < y < 1\}$ *where* $\phi^+(x, y) = y$
- $\phi^- = \{(x, y) \in S^1 \mid x < 0\} \to \{y \mid -1 < y < 1\}$ *where* $\phi^-(x, y) = y$
- $\psi^+ : \{(x, y) \in S^1 \mid y > 0\} \to \{x \mid -1 < x < 1\}$ *where* $\psi^+(x, y) = x$
- $\psi^- : \{(x, y) \in S^1 \mid y < 0\} \to \{x \mid -1 < x < 1\}$ *where* $\psi^-(x, y) = x$

So, for example, $\psi^+ \circ (\phi^+)^{-1} : \{y \mid 0 < y < 1\} \to \{x \mid 0 < x < 1\}$ *where* $\psi^+ \circ (\phi^+)^{-1}(y) = \psi^+(\sqrt{1 - y^2}, y) = \sqrt{1 - y^2}$.

Notice that $\psi^+ \circ (\phi^+)^{-1}(y) = \sqrt{1 - y^2}$ is in fact smooth for all $0 < y < 1$. So far we have only tested one of the overlaps for smoothness, to finish we must also check the rest of the overlaps!

Before we find a better way to check whether or not something is a manifold, let us observe this process again with a torus.

Example 3.15. *The Torus is a Manifold. Consider the set defined by*

$$T = \left\{(x, y, z) \;\middle|\; (3 - \sqrt{x^2 + y^2})^2 + z^2 = 1\right\}$$

This is a torus in $\mathbb{R}^3$. If $f : \mathbb{R}^3 \to \mathbb{R}$ is defined by $f(x, y, z) = (3 - \sqrt{x^2 + y^2})^2 + z^2$, then $T = f^{-1}(\{1\})$ is a level set. With a proper restriction to u and v, the map

$$\tau^{-1}(u, v) = \langle (3 + \cos(u)) \cos(v), (3 + \cos(u)) \sin(v), \sin(u) \rangle$$

is a coordinate patch, meaning τ is a coordinate chart on T. In fact, by constructing several τ's with different restrictions on u and v, to ensure τ is invertible, the entirety of T can be covered, yielding an atlas.

Let $\psi : \left\{(x,y,z) \in T \;\middle|\; z > 0\right\} \to \left\{(x,y) \;\middle|\; 4 < x^2 + y^2 < 16\right\}$ be defined by $\psi(x,y,z) = (x,y)$ To confirm that ψ is invertible and a homeomorphism, we first find that

$$\psi^{-1}(x,y) = \left(x, y, \sqrt{1 - (3 - \sqrt{x^2 + y^2})^2}\right)$$

For the specific choice of domain for τ^{-1} given as $(u,v) \in (0, 2\pi) \times (-\frac{\pi}{2}, \frac{\pi}{2})$, we find

$$\psi \circ \tau^{-1} : (0,\pi) \times (-\pi/2, \pi/2) \to \left\{(x,y) \;\middle|\; 4 < x^2 + y^2 < 16 \textit{ and } x > 0\right\}$$

is given by $\psi \circ \tau^{-1}(u,v) = ((3 + \cos(u))\cos(v), (3 + \cos(u))\sin(v))$. We need to see that this is a diffeomorphism to get that ψ and τ are compatible. If fact, to be complete, the seven other transition functions must be checked as well. For now we will simply show that ψ is compatible with T's atlas without worrying for specific domains for the various τ's. Again, observe that

$$\psi \circ \tau^{-1}(u,v) = \psi(\tau^{-1}(u,v)) = \psi(3 + \cos(u))\cos(v), (3 + \cos(u))\sin(v), \sin(u))$$

$$= ((3 + \cos(u))\cos(v), (3 + \cos(u))\sin(v))$$

The Jacobian matrix of $\psi \circ \tau^{-1}$ is calculated as

$$J_{\psi\circ\tau^{-1}} = \begin{bmatrix} -\sin(u)\cos(v) & -\sin(v)(3 + \cos(u)) \\ -\sin(u)\sin(v) & \cos(v)(3 + \cos(u)) \end{bmatrix}.$$

where

$$\det(J_{\psi\circ\tau^{-1}}) = -\sin(u)(3 + \cos(u)) \neq 0$$

for all u and v in the domain (recall $0 < u < \pi$ so that $\sin(u) > 0$). If we checked the remaining transition functions, we would have shown that T is indeed a 2-manifold.

It should be quite apparent from the last example that testing whether something is a manifold can become computationally impractical very quickly. To resolve this we need a little more machinery.

Definition 3.16. *Let $\mathcal{M}$ and $\mathcal{N}$ be smooth manifolds of dimension m and n respectively. A mapping $f : \mathcal{M} \to \mathcal{N}$ is a smooth map at $p \in \mathcal{M}$ if and only if there is a chart on $\mathcal{M}$, (U, ϕ) with $p \in U$ and a chart on $\mathcal{N}$, (V, ψ) with $f(p) \in V$, such that*

$$\psi \circ f \circ \phi^{-1} : \phi(U \cap f^{-1}(V)) \to \psi(f(U) \cap V)$$

is smooth at p. Specifically, if $f(x^1, ..., x^m) = (y^1, ..., y^n)$ is f expressed in coordinates on U and V, then smoothness at p means $y^i(x^1, ..., x^m)$ is smooth at p for every i.

It must be said that we are employing a serious abuse of notation here. The mapping $f(x^1, ..., x^m) = (y^1, ..., y^n)$ actually refers to the composition $\psi \circ f \circ \phi^{-1}$, as unfortunate as it may seem, doing this is very useful and is a standard practice. Regardless, if f is a smooth real valued function on $\mathcal{M}$ then we write $f \in \Omega^0(\mathcal{M}) = \{f : \mathcal{M} \to \mathbb{R} \mid f \text{ is smooth}\}$. It is interesting to note here that $\Omega^0(\mathcal{M})$ forms an algebra, however we will not discuss this in detail here.

Definition 3.17. *If a mapping $f : \mathcal{M} \to \mathcal{N}$ is a homeomorphism and both f and f^{-1} are smooth, then we say f is a diffeomorphism.*

Let $\mathcal{M}$ and $\mathcal{N}$ be smooth manifolds where $\dim(\mathcal{M}) = m$ and $\dim(\mathcal{N}) = n$ and consider a smooth mapping $f : \mathcal{M} \to \mathcal{N}$ that is defined by $f(x^1, ..., x^m) = (y^1(x^1, ..., x^n), ..., y^n(x^1, ..., x^n))$. If $m \leq n$ and the Jacobian Matrix $\left(\frac{\partial y^i}{\partial x^j}\right)$ has maximal rank $= m$, at $p \in \mathcal{M}$ then f is called an immersion at p. In the case of f being an immersion for all points $p \in \mathcal{M}$, then $\mathcal{M}$ is an immersed submanifold of $\mathcal{N}$. If $m \geq n$ and and the Jacobian matrix has maximal rank $= n$ at $p \in \mathcal{M}$, then we say f is a submersion at p. We say an injective immersion is called an embedding, or imbedding, provided that f maps $\mathcal{M}$ homeomorphically onto its image $f(\mathcal{M})$, in the induced topology.

The difference between immersions and embeddings is that the image of an immersion can self-intersect whereas the image of an embedding cannot. This implies that if something is an immersion, then it need not be a manifold, see the figure eight as an example.

Theorem 3.18. *Whitney Embedding Theorem. Given a manifold $\mathcal{M}$ where $\dim(\mathcal{M}) = n$, then*

1. *If $\mathcal{M}$ is a topological manifold then $\mathcal{M}$ can be embedded in $\mathbb{R}^{2n+1}$.*

2. *If $\mathcal{M}$ is a smooth manifold then $\mathcal{M}$ can be embedded in $\mathbb{R}^{2n}$.*

See, for example, [Sk].

Theorem 3.19. *The Inverse Function Theorem. Let U be an open subset of $\mathbb{R}^n$ and define a smooth function f defined as*

$$f : U \to \mathbb{R}^n.$$

Let $a \in U$ and let $J_f(a) = Df(a)$ (the Jacobian matrix) be invertible. Then there exists some open neighborhood of a, $V \subseteq U$, such that

$$f : V \to f(V)$$

is a diffeomorphism. Moreover, we have

$$Df^{-1}(f(a)) = (Df(a))^{-1}.$$

Note that if $f : V \to f(V)$ is a diffeomorphism, then $(Df^{-1}(f(a))) \circ Df(a) = D(f^{-1} \circ f)(a) = D(\mathrm{Id_V})(a) = I_n$ by the chain rule. Thus $Df^{-1}(f(a)) = (Df(a))^{-1}$, so the converse is true as well. Again consider the map $f : \mathcal{M} \to \mathcal{N}$ with $m \geq n$. If f is a submersion at $p \in \mathcal{M}$, say $f(p) = q$ then p is called a regular point of f, otherwise p is a critical point of f. A point $q \in \mathcal{N}$ is a regular value of f if every point in $f^{-1}(q)$ is regular. It is also good to note, $f^{-1}(q)$ in general is called the fiber over q.

Theorem 3.20. *Regular Value Theorem. Let $f : \mathcal{M} \to \mathcal{N}$ be a smooth map between manifolds with $\dim(\mathcal{M}) = m$ and $\dim(\mathcal{N}) = n$ where $m \geq n$. In addition, let $q \in \mathcal{N}$ be a regular value of f. Then $f^{-1}(q)$ is a smooth embedded sub-manifold of $\mathcal{M}$ with $\dim(f^{-1}(q)) = m - n$.*

The proof of this theorem is given in [R]. The results of this theorem are very powerful. Now to show something is a manifold, simply take a smooth map between two known manifolds and show the Jacobian matrix is non singular at the appropriate points.

Example 3.21. *Revisiting the circle: Let* $f : \mathbb{R}^2 \to \mathbb{R}$ *defined by* $f(x, y) = x^2 + y^2$ *and observe the level set at* $f(x, y) = 1$*. Then the Jacobian matrix is* $J_f = [f_x\ f_y] = [2x\ 2y]$*.*
Consider $S^1 = f^{-1}(\{1\}) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$*. Notice that* f *is in fact a smooth map and that* J_f *is not zero at any point in* S^1*, so the rank of* J_f *is 1 (i.e., maximal) at every point in* S^1*. By the Regular Value Theorem,* S^1 *is an embedded submanifold of* $\mathbb{R}^2$ *of dimension* $\dim(S^1) = 2 - 1 = 1$*.*

Example 3.22. *Revisiting a Torus as a Manifold: Again consider*

$$T = \left\{ (x, y, z) \;\middle|\; (3 - \sqrt{x^2 + y^2})^2 + z^2 = 1 \right\},$$

a torus in $\mathbb{R}^2$*. Notice if we define* $f : \mathbb{R}^3 \to \mathbb{R}$ *by*

$$f(x, y, z) = (3 - \sqrt{x^2 + y^2})^2 + z^2.$$

Once again, $T = f^{-1}(\{1\})$ *is a level set. We calculate the Jacobian*

$$J_f = [f_x\ f_y\ f_z] = \begin{bmatrix} 2x - \frac{-6x}{\sqrt{x^2+y^2}} & 2y - \frac{-6y}{\sqrt{x^2+y^2}} & 2z \end{bmatrix}$$

Note that f *is smooth everywhere except at* $(0, 0, z)$ *and the Jacobian has max rank if we stay away from these problem points. However, each of these points are not in* $f^{-1}(\{1\})$*. Therefore,* $f^{-1}(1) = T$ *is a* $3 - 1 = 2$*-manifold.*

This approach is clearly more computationally practical than examples 3.14 and 3.15. Recall that using the first method would have required checking overlap mappings for pairs of our four charts for the circle and eight charts for the torus.
As a final note on manifolds, for this paper any manifold discussed will be of finite dimension. There are various theories of infinite dimensional manifolds, they can look quite different as many things do not hold true for infinite dimensional spaces.

4 Tangential Structures

Now that we have constructed the idea of a manifold, we must develop the notion of what is tangent to one. As it turns out, there is more structure to the tangent spaces of a manifold than the vector space structure alone. The following introduction to tangential structures has been compiled from [R] chapter three and [Sp] chapter three.

4.1 Tangent Spaces on $\mathcal{M}$

A tangent space at a point p, on a manifold $\mathcal{M}$, is denoted as $T_p\mathcal{M}$. The reader should be very familiar already with tangent planes as seen in the figure below. Our goal here is to generalize to more abstract spaces that cannot always be visualized.

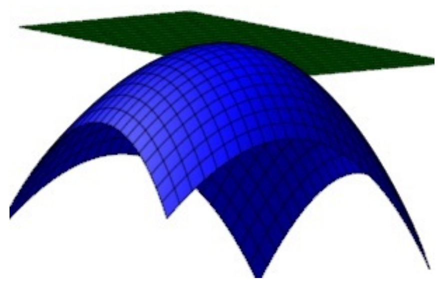


Figure 2: Tangent space at a point on a surface.

There are three main viewpoints to the tangent space of a given manifold. First, we will note that tangent vectors at a point $p \in \mathcal{M}$ can be considered as equivalence classes of curves through p whose tangent information matches. This is "geometric" formulation however it will not be the approach we take. Next, we have the coordinate approach.

Theorem 4.1. *Consider the local coordinates $x^1, ..., x^n$ in a neighborhood U of a point $p \in \mathcal{M}$. Then $T_p\mathcal{M}$ is spanned by the tangent vectors $\frac{\partial}{\partial x^1}\Big|_p, ..., \frac{\partial}{\partial x^n}\Big|_p$ and $\dim(T_p\mathcal{M}) = n$.*

$Proof$ See [R].

These derivatives of coordinate functions, $\partial/\partial x^i\Big|_p$, at a point give us a computational viewpoint and is more of a physics friendly viewpoint.

These two viewpoints are both worth mentioning, however the definition we will take is easier to use in proofs and generalizations. Note that it would be an interesting exercise for the reader to show all three are equivalent. First, recall that $\Omega^0(\mathcal{M}) = \{h : \mathcal{M} \to \mathrm{R} \mid h \text{ is smooth}\}$ is the (algebra) of smooth scalar valued functions defined on a manifold $\mathcal{M}$.

Definition 4.2. *Given a manifold $\mathcal{M}$, a tangent vector X_p at a point $p \in \mathcal{M}$ is a linear derivation at p. That is to say, for any scalars $\alpha, \beta \in \mathbb{R}$ and scalar valued functions $f, g \in \Omega^0(\mathcal{M})$ the following conditions must be satisfied for $X_p : \Omega^0(\mathcal{M}) \to \mathbb{R}$*

1. $X_p(\alpha f + \beta g) = \alpha X_p(f) + \beta X_p(g)$ *(X_p is $\mathbb{R}$-linear),*

2. $X_p(fg) = g(p) X_p(f) + f(p) X_p(g)$ *(the Leibniz property).*

Given some coordinate chart ϕ, it is easy to see that $\left.\frac{\partial}{\partial x^i}\right|_p : \Omega^0(\mathcal{M}) \to \mathbb{R}$ are linear derivations at p where given $f \in \Omega^0(\mathcal{M})$ we have $(f \circ \phi^{-1})(x^1, \ldots, x^n)$ is real valued and then $\left.\frac{\partial}{\partial x^i}\right|_p (f) = \left.\frac{\partial (f \circ \phi^{-1})}{\partial x^i}\right|_p$. By now it should be apparent that we must visualize tangents in a new way, one that uses the tangent vector as a tool that tells us how a function changes as we move in a given direction. As an example for why it makes sense to view derivatives as vectors, let ψ be a scalar field on $\mathbb{R}^2$ (i.e., a real valued function defined on $\mathbb{R}^2$) and let $\phi : I \to \mathbb{R}^2$ be a parameterized curve denoted as $s \mapsto (\phi^1(s), \phi^2(s)) = (x^1(s), x^2(s))$. The derivative gives, via the chain rule

$$\frac{d}{ds}\psi(\phi(s)) = \frac{d\psi}{dx^1}\frac{dx^1}{ds} + \frac{d\psi}{dx^2}\frac{dx^2}{ds} = \frac{d\phi}{ds} \cdot \nabla\psi \tag{4.1}$$

If we remove the arbitrary ψ from the expression, we see the intuitive nature of the differential as

$$\frac{d}{ds} = \frac{d\phi}{ds} \cdot \nabla \tag{4.2}$$

This tells us that the differential acts like a tangent vector to a curve however, this notion does not make sense in the absence of one. This concept extends naturally to our definition of the tangent space for some smooth map f. First consider a manifold $\mathcal{M}$ of dimension n with a coordinate chart (U, ϕ) where $p \in \mathcal{M}$ and fix

$$\phi(p) = (x^1(p), ..., x^n(p)) = \boldsymbol{a} = (a^1, ..., a^n). \tag{4.3}$$

Define a smooth mapping

$$f : \phi(U) \to \mathbb{R} \tag{4.4}$$

such that

$$f \circ \phi^{-1} = f(x^1, ..., x^n). \tag{4.5}$$

Observe the expansion, via Taylor's theorem, of f in the neighborhood of p with local coordinates

$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + f_{x^1}(\mathbf{a})(x^1 - a^1) + ... + f_{x^n}(\mathbf{a})(x^n - a^n) + R(\boldsymbol{x})((x^1 - a^1)^2 + ... + (x^n - a^n)^2). \tag{4.6}$$

Then by applying X_p to f we find

$$X_p(f) = \sum_i f_{x^i}(\mathbf{a}) X_p(x^i - a^i) = \sum_i f_{x^i}(\mathbf{a}) X_p(x^i) \tag{4.7}$$

Note that $\mathbf{a} = \phi(p)$ so actually $f(\mathbf{a}) = f(p)$ via our standard abuse of notation. From here on it will be of great benefit for all involved to employ the Einstein summation notation where 4.7 can be realized as

$$X_p(f) = f_{x^i}(p) X_p(x^i) = X_p(x^i) \frac{\partial f}{\partial x^i}\bigg|_p$$

Again after removing our test function, we have the result

$$X_p = X_p(x^i) \frac{\partial}{\partial x^i}\bigg|_p \tag{4.8}$$

If the $X^i(p)$ are scalar valued components of X_p

$$X_p = X^i(p) \frac{\partial}{\partial x^i}\bigg|_p \tag{4.9}$$

then $X^i(p) = X_p(x^i)$.

Removing our evaluation at p and requiring the X^i's to vary smoothly over the entire patch, or even over all of $\mathcal{M}$, we have the definition of a vector field X on $\mathcal{M}$, which is a smooth map $p \mapsto X_p$. In terms of local coordinates we have

$$X = X^i \frac{\partial}{\partial x^i} = X(x^i)\frac{\partial}{\partial x^i}. \tag{4.10}$$

Another way to look at this is, a vector field X is a smooth derivation on the algebra $\Omega^0(\mathcal{M})$. In particular, $X(\alpha f) = \alpha X(f)$, $X(f+g) = X(f) + X(g)$, and $X(fg) = X(f)g + fX(g)$ for all $\alpha \in \mathbb{R}$, $f, g \in \Omega^0(\mathcal{M})$. In this paper, we will denote the set of all vector fields on a manifold $\mathcal{M}$ as $\mathfrak{X}(\mathcal{M})$. A vector field on $\mathcal{M}$ is an example of a section on the tangent bundle of $\mathcal{M}$, two concepts we will begin to explore in detail later.

To sum up some of the new notation from this section, recall that for an n-manifold $\mathcal{M}$ with coordinates $\{x^i\}, \ \{y^j\}$ for the charts ϕ and ψ respectively,

$$X_p = X^i_p \frac{\partial}{\partial x^i}\Big|_p, \qquad X^i_p = X^p(x^i), \tag{4.11}$$

$$Y_p = Y^j_p \frac{\partial}{\partial y^j}\Big|_p, \qquad Y^j_p = Y^p(y^j). \tag{4.12}$$

We need to now pin down how differential interactions between two charts unfold. Since the span of $T_p\mathcal{M}$ is $\{\frac{\partial}{\partial x^i}\}$ it follows from the chain rule that

$$\frac{\partial}{\partial y^j} = \frac{\partial x^i}{\partial y^j}\frac{\partial}{\partial x^i} \tag{4.13}$$

therefore by expanding in the basis given in 4.12 we find that the tangent vectors on $\mathcal{M}$ can be expressed in terms of the coordinates corresponding to ψ as

$$X_p = X_p(x^i)\frac{\partial}{\partial x^i}\Big|_p = X_p(x^i)\frac{\partial y^j}{\partial x^i}\frac{\partial}{\partial y^j}\Big|_p = X_p(y^j)\frac{\partial}{\partial y^j}\Big|_p. \tag{4.14}$$

This result should not be entirely too surprising. In fact, the result given in 4.12 is simply the change of basis matrix given as the inverse of the Jacobian matrix.

4.2 Cotangent Space

Recall that for a vector space V, there exists a dual space V^*. Using the tangent space $T_p\mathcal{M}$ of a manifold $\mathcal{M}$ as an example, we can analyze its dual, $(T_p\mathcal{M})^*$, denoted also as $T_p^*\mathcal{M}$. This dual space is called the cotangent space of $\mathcal{M}$ at a point p whose elements are called cotangent vectors at p. As with general vector spaces, these covectors are linear maps from $T_p^*\mathcal{M}$ to $\mathbb{R}$.

Definition 4.3. *A* 1*-form on a manifold $\mathcal{M}$ is a function that assigns points on $\mathcal{M}$ to an element $\theta_p \in T_p^*\mathcal{M}$ where*

$$\theta_p : T_p\mathcal{M} \to \mathbb{R}$$

is linear and this choice varies smoothly with p. Denote the space of all 1-forms on $\mathcal{M}$ by $\Omega^1(\mathcal{M}) = \mathfrak{X}^*(\mathcal{M})$.

Let $X \in \mathfrak{X}(\mathcal{M})$, then assign each $p \in \mathcal{M}$ the value θX is defined as $\theta_p(X_p)$. If θX is smooth for all $X \in \mathfrak{X}(\mathcal{M})$, then θ is smooth. If we consider $f \in \Omega^0(\mathcal{M})$, the differential of f, denoted df, is the 1-form such that

$$(df)(X_p) = X_p(f) \tag{4.15}$$

for every $X_p \in T_p\mathcal{M}$. Note that

$$(df)\Big|_p : T_p\mathcal{M} \to \mathbb{R} \tag{4.16}$$

is linear and

$$(df)(X) = X(f) \tag{4.17}$$

is smooth. These differentials satisfy the following three properties:

1. $d : \Omega^0(\mathcal{M}) \to \Omega^1(\mathcal{M}) = \mathfrak{X}^*(\mathcal{M})$ is $\mathbb{R}$ linear.
2. If $f, g \in \Omega^0(\mathcal{M})$, then $d(fg) = gdf + fdg$.
3. If $f \in \Omega^0(\mathcal{M})$ and $g \in \Omega^0(\mathbb{R}) = \{\text{smooth functions on } \mathbb{R}\}$ then $d(h(f)) = h'(f)df$.

If $\{x^i\}$ are local coordinates for an open subset $U \subset \mathcal{M}$ then the differential of each coordinate gives $\{dx^i\}$, known as the set of coordinate differentials. Note that $\{dx^i\}$ is the dual basis to the coordinate vector fields $\{\partial/\partial x^j = \partial_j\}$ since

$$dx^i(\partial_j) = \frac{\partial x^i}{\partial x^j} = \delta^i_j. \tag{4.18}$$

Let $\{v_i\}$ be a basis for V and $\{w^i\}$ the dual basis for V^*. Then for a general $h \in V^*$ expand $h = h_i w^i$. It follows that

$$\begin{aligned} h(v_j) &= h_i w^i(v_j) \\ &= h_i \delta^i_j \\ &= \ h_j \end{aligned}$$

Notice how $h(v_j)$ simply selects h's j-th component. It follows that for a general 1-form θ

$$\theta = \theta(\partial_i) dx^i \ \text{ on } \ U \subseteq \mathcal{M} \tag{4.19}$$

where U our chart's domain. If we revisit $f \in \Omega^0(\mathcal{M})$, then

$$df(\partial_i) = \frac{\partial f}{\partial x^i} \tag{4.20}$$

from which we conclude

$$df = \frac{\partial f}{\partial x^i} dx^i. \tag{4.21}$$

Since a differential 1-form, ω is a smooth assignment of $p \mapsto \omega_p$, ω is a smooth covector field on $\mathcal{M}$. For local coordinates on the manifold, a general 1-form is expressed as

$$\omega = \omega_i dx^i = \omega(\partial/\partial x^i) dx^i. \tag{4.22}$$

The notion of a differential k-form is an extension of this idea where, if η is a k-form then $\eta \in \bigwedge^k T_p\mathcal{M}$. Differential k-forms are crucial to integration on manifolds, among other things, yet not important for our purposes. For more information see [R].

4.3 Tangent Bundles

Now that we have an idea of what it entails to be a tangent space to a manifold, we may use these ideas to construct more interesting mathematical objects. As an example, we begin with a special case of the more general vector bundle which will be covered in detail later. To build intuition, consider $\mathcal{M} = \mathbb{R}^n$. To denote a given tangent vector at a point p we use (p, v), the set of all such pairs is simply $\mathbb{R}^n \times \mathbb{R}^n$. This is known as the tangent bundle of $\mathbb{R}^n$ and is denoted $T\mathbb{R}^n$. In general, take the following definition.

Definition 4.4. *Given a smooth manifold $\mathcal{M}$, with a tangent space $T_p\mathcal{M}$ for each $p \in \mathcal{M}$, we construct the tangent bundle $T\mathcal{M}$ which joins together all tangent vectors in $\mathcal{M}$ as the disjoint union of all tangent spaces,*

$$T\mathcal{M} = \dot{\bigsqcup_{p\in\mathcal{M}}} T_p\mathcal{M} = \bigcup_{p\in\mathcal{M}} \{p\} \times T_p\mathcal{M}$$

For this we construct the bundle projection map defined as

$$\pi : T\mathcal{M} \to \mathcal{M} \tag{4.23}$$

where,

$$\pi(p, v) = p. \tag{4.24}$$

Given a vector in the tangent bundle, the projection map tells you where you are. The set $\pi^{-1}(p) = \{p\} \times T_p\mathcal{M}$ gives all tangent vectors located at p also known as the fiber over p. Each fiber can be turned into a vector space if we define

$$(p, v) \oplus (p, w) = (p, v + w)$$
$$a \cdot (p, v) = (p, a \cdot v).$$

Note that the operations $\oplus$ and $\cdot$ are defined over $\bigcup_{p\in\mathcal{M}} \pi^{-1}(p) \times \pi^{-1}(p)$ and $\mathbb{R} \times T\mathcal{M}$ respectively. Often it is convenient to consider the tangent bundle as a four-tuple $(T\mathcal{M}, \mathcal{M}, \mathbb{R}^n, \pi)$. Here we say $T\mathcal{M}$ is the total or bundle space, $\mathcal{M}$ is the base space, $\mathbb{R}^n$ is our fiber (vector) space, and π is again the projection map.

Example 4.5. *Consider* $(T\mathcal{M}, \mathcal{M}, \mathbb{R}^n, \pi)$ *where* $\pi(p, v) = p$ *and* $v = (p, v) \in T_p\mathcal{M}$. *In standard coordinates for* $\mathbb{R}^n$ *we have* $x^1, ..., x^n$ *and* $v = v^i \frac{\partial}{\partial x^i}$. *Observe that* $T\mathcal{M}$ *has coordinates* $(p^1, \ldots, p^n, v^1 \ldots, v^n)$ *where* $(x^1, \ldots, x^n)(p) = (x^1(p), \ldots, x^n(p)) = (p^1, \ldots, p^n)$*, yielding* $\dim(T\mathcal{M}) = 2n$. *We are applying* $(x^1, \ldots, x^n, dx^1, \ldots, dx^n)$ *to* (p, v). *So given a chart* $\varphi : U \to \mathbb{R}^n$ *on* $\mathcal{M}$ *where* $\varphi(p) = (x^1(p), \ldots, x^n(p)) = (p^1, \ldots, p^n)$ *then* $T\mathcal{M}$ *has a chart:* $\psi : TU \to \mathbb{R}^{2n}$ *where* $p \in U$, $(p, v) \in T_p\mathcal{M}$*, and* $\psi(p, v) = (x^1(p), \ldots, x^n(p), dx^1(v), \ldots, dx^n(v)) = (p^1, \ldots, p^n, v^1, \ldots, v^n)$. *So briefly,* $\varphi = (x^1, \ldots, x^n)$ *and* $\psi = (x^1, \ldots, x^n, dx^1, \ldots, dx^n)$

Theorem 4.6. *By equipping* $T\mathcal{M}$ *with the standard topology and a smooth atlas induced from the atlas on* $\mathcal{M}$ *(as defined above), we have that the tangent bundle to a manifold is itself a manifold and if* $\dim(\mathcal{M}) = n$ *then we have* $\dim(T\mathcal{M}) = 2n$.

Sketch of proof: Let (U, ϕ) be a chart on M, then we get a chart $(TU, \hat{\phi})$ on $T\mathcal{M}$ as follows:

$$\hat{\phi}(p, X_p) = (x^1(p), \ldots, x^n(p), X_p^1, \ldots X_p^n)$$

where

$$\phi(p) = (x^1(p), \ldots, x^n(p)) \text{ and } X_p = X_p^i \left.\frac{\partial}{\partial x^i}\right|_p.$$

Notice that $\hat{\phi}(p, X_p)$ is smooth since each of it's components are themselves smooth and that there are $2n$ components to the inputs.

It is interesting to note, that since TM is itself a manifold, we can construct higher order tangent bundles where,

$$T^2M = T(TM) \tag{4.25}$$

and in general this can be defined recursively as,

$$T^jM = T(T^{j-1}M) \tag{4.26}$$

Where the $\dim(T^jM) = (2^j \dim(\mathcal{M}))$.

It is clear to see by our discussion above, each of these higher order tangent bundles are themselves manifolds. A simple example to visualize a first order bundle can be expressed with a circle where the tangent lines are joined in a smooth non-overlapping fashion,

As can be seen, the dimension of this tangent bundle is twice that of the circle. With this new notion of a tangent bundle, we can view objects such as vector fields in a new light.

Definition 4.7. *A (smooth) vector field X along a smooth map $\psi : U \subset \mathcal{M} \to \mathcal{M}$ is a mapping $X : U \to T\mathcal{M}$ such that $\pi \circ X = \psi$. Moreover, a (smooth) vector field X on $\mathcal{M}$ is a smooth map $X : \mathcal{M} \to T\mathcal{M}$ such that $\pi \circ X = id_{\mathcal{M}}$.*

That is to say, each point $p \in \mathcal{M}$ is assigned a tangent vector to $\mathcal{M}$ by X at p. This view gives a great insight into the applications of vector fields in the physical sciences for example, if α is a curve in $\mathcal{M}$ then it's velocity α' is a vector field on α. From here on when vector fields are mentioned, they will be taken to be smooth unless otherwise noted. The same will hold for a covector field, which will be defined in a similar way shortly.

4.4 Cotangent Bundle

The cotangent bundle is defined just as one might expect considering the definition of the tangent bundle. Recall that given a point p within the coordinate patch $U \subset \mathcal{M}$, expressed in local coordinates $x^1, \ldots, x^n$ and the collection $dx^1, \ldots, dx^n$ (evaluated at p) form a basis of the cotangent space $T_p^*\mathcal{M}$. In particular, $\alpha \in T_p^*\mathcal{M}$ can be expressed as $\alpha = \alpha_i(p)dx^i$.

Definition 4.8. *A cotangent bundle to the n-manifold $\mathcal{M}$ is the space $T^*\mathcal{M}$ of all covectors located at each point on $\mathcal{M}$. Like with the tangent bundle, a point in $T^*\mathcal{M}$ is a pair (p, α) where α is a covector located at $p \in \mathcal{M}$. That is*

$$T^*\mathcal{M} = \dot{\bigsqcup_{p\in\mathcal{M}}} T_p^*\mathcal{M} = \bigcup_{p\in\mathcal{M}} \{p\} \times T_p^*\mathcal{M}$$

As before we need a projection mapping, call it π where analogous to before

$$\pi : T^*\mathcal{M} \to \mathcal{M} \tag{4.27}$$

is defined by $(p, \alpha) \mapsto p$ and thus

$$\pi^{-1}(p, \alpha) = \{p\} \times T_p^*\mathcal{M}. \tag{4.28}$$

The proof that the cotangent bundle is also a manifold of $\dim(T^*\mathcal{M}) = 2n$ is quite similar to that for the tangent bundle and will be omitted from this discussion, but should seem very plausible to the reader at this point.

Definition 4.9. *A covector field Z on a smooth map $\psi : U \subset \mathcal{M} \to \mathcal{M}$ is a mapping $Z : U \to T^*\mathcal{M}$ such that $\pi \circ Z = \psi$. Moreover, a covector field Z on $\mathcal{M}$ is a smooth map $Z : \mathcal{M} \to T^*\mathcal{M}$ such that $\pi \circ Z = id_{\mathcal{M}}$.*

This notion of a cotangent bundle has many applications in physics, for example the phase space for Hamiltonian Dynamics. For further details see [F] page 54. Notice that the structure of the cotangent bundle is almost the same as that of the tangent bundle. This hints to them being special cases of a more general underlying structure, as it turns out this is true and we will explore this generalization further.

4.5 Vector Bundles

With tangent and cotangent bundles for $\mathcal{M}$ in mind as special cases, we generalize one final time to what is called a vector bundle. Though we will not explore the idea in detail, we can extend these bundles to general tensor fields as well. So, for the purpose of this paper we will only consider vector spaces and their dual.

Definition 4.10. *A vector bundle is the 4-tuple $(E, \mathcal{M}, Y, \pi)$ where E is the total space or bundle space, $\mathcal{M}$ is the base space, Y is a vector space, and π is the projection mapping. Note that we will generally consider cases where Y is $\mathbb{R}^m$.*

In general if $\dim(E) = m$ then m is the rank of the bundle. As a short hand, we will call upon $(E, \mathcal{M}, Y, \pi)$ as simply E. The following three conditions for E are also required:

1. Each fiber $\pi^{-1}(p)$, over a point $p \in \mathcal{M}$ is isomorphic to a fixed vector space Y,

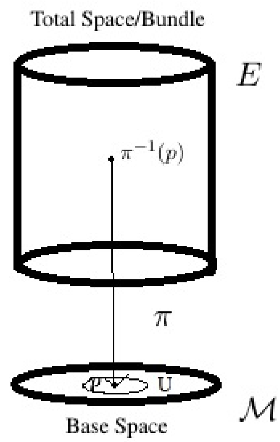


Figure 3: Depiction of a tangent bundle of a manifold $\mathcal{M}$.

$$\pi^{-1}(p) = \{x \in E \mid \pi(x) = p\} \cong Y \cong \mathbb{R}^m.$$

2. E is locally trivializable. Meaning for all $p \in \mathcal{M}$ there exists an open neighborhood of p, say $U \subseteq \mathcal{M}$ and a diffeomorphism ϕ_U such that

$$\phi_U : \pi^{-1}(U) \to U \times Y.$$

3. The local trivialization ϕ_U carries fibers to fibers and its restriction to a fiber is linear. That is

$$\phi_U \Big|_{\pi^{-1}(p)} : \pi^{-1}(p) \to \pi_1^{-1}(p) = \{p\} \times Y,$$

is a (linear) isomorphism where for $y \in Y$

$$\pi_1 : U \times Y \to U \quad \text{is defined by} \quad \pi_1(p, y) = p.$$

Example 4.11. *The tangent bundle is a vector bundle with the 4-tuple $(T\mathcal{M}, \mathcal{M}, \mathbb{R}^n, \pi)$ where $\pi(p, v) = p$ and $v = (p, v) \in T_p\mathcal{M}$. Just as easily, we can see the cotangent bundle is a vector bundle with $(T^*\mathcal{M}, \mathcal{M}, \mathbb{R}^n, \pi)$.*

Example 4.12. *Every manifold $\mathcal{M}$ is a rank 0 vector bundle as follows: let $E = \mathcal{M}$, $\mathcal{M} = \mathcal{M}$, $V = \{0\}$, and $\pi = id_{\mathcal{M}}$.*

4.6 Sections

Recall that the projection mapping on a bundle gives a mapping from the bundle to the base manifold. To move from the manifold up to the bundle, we define what is called a section. Given the total space E of a vector bundle $(E, \mathcal{M}, Y, \pi)$, a smooth section, call it $\mathfrak{s}$, of E is a smooth map

$$\mathfrak{s} : \mathcal{M} \to E \tag{4.29}$$

such that for each $p \in \mathcal{M}$ we select some element of the fiber $\mathfrak{s}(p) \in \pi^{-1}(p)$. This means that

$$\pi \circ \mathfrak{s} = id_{\mathcal{M}}. \tag{4.30}$$

We say $\mathfrak{s}$ is a section and denote the space of sections on a bundle E as $\Gamma(E)$. Since $\mathfrak{s}$ can be viewed as a vector valued mapping, a vector space structure is induced on $\Gamma(E)$.

Example 4.13. *As a quick example, $\mathfrak{X}(\mathcal{M}) = \Gamma\mathcal{M}$, which as we know are vector fields on the manifold $\mathcal{M}$. Dual to this notion, $\mathfrak{X}^*(\mathcal{M}) = \Gamma^*\mathcal{M}$, covector fields are sections of the cotangent bundle. More generally, sections of tensor bundles yield tensor fields.*

Just as $T\mathcal{M}$ and $T^*\mathcal{M}$ are built from $T_p\mathcal{M}$ and its dual, one can build bundles from tensor products of $T_p\mathcal{M}$ and it dual to get tensor bundles. Sections of these bundles are known as tensor fields.

There is much more to learn about both general bundles and sections of them, this taste should give the reader just enough to move forward with a sense of what sort of structures one might be working with. These generalizations help to tie together several ideas previously introduced and show that they are all more or less instances of the same general idea. For now we will move into some differential geometry.

5 Semi-Riemannian Geometry

It is possible that the reader has heard of Riemannian geometry at some point in their studies. The focus of this section will however be on pseudo-Riemannian manifolds, known also as semi-Riemannian manifolds. We will discuss the details for what each of these mean soon, but it is good to note that our choice for the more general semi-Riemannian geometry allows for a structure for which general relativity can be developed. The following developments in semi-Riemannian geometry have been constructed from material found in [LEE] chapters twelve and thirteen in addition to [ON] chapters two and three.

5.1 Lie Algebras and Derivatives

Before we break into semi-Riemannian geometry we first need to briefly develop a very important mathematical concept, namely Lie algebras. For a more detailed introduction to Lie algebras see [EW].

Definition 5.1. *For a given field $\mathbb{F}$, a Lie algebra over $\mathbb{F}$ is a vector space V over $\mathbb{F}$, equipped with a bilinear map, called the Lie bracket*

$$[\cdot,\cdot] : V \times V \to V \tag{5.1}$$

where for $x, y \in V$, $(x, y) \mapsto [x, y]$; and the following properties are satisfied:

1. $[x, x] = 0$ *for all $x \in V$,*

2. $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ *for all $x, y, z \in V$ (Jacobi identity).*

In the context of an associative algebra, the bracket $[x, y]$ can be defined to be the commutator of x and y: $[x, y] = xy - yx$.

Example 5.2. *Consider the set of all linear maps $T : V \to V$ denoted as $gl(V)$. It can be shown that $gl(V)$ is again a vector space over $\mathbb{F}$ and hence we can define a Lie bracket by*

$$[x, y] = x \circ y - y \circ x \quad \text{for all } x, y \in gl(V).$$

Here $\circ$ is the composition of maps. This example highlights the name "commutator" as the bracket is essentially a measurement of how much the operation fails to commute.

Definition 5.3. *Given $X, Y \in \mathfrak{X}(\mathcal{M})$, the Lie derivative with respect to X is defined as the map*

$$\mathcal{L}_X Y := [X, Y].$$

Since we are dealing with a Lie algebra structure, the bracket must satisfy the conditions outlined above, therefore from the Jacobi identity and for any $X, Y, Z \in \mathfrak{X}(M)$

1. $\mathcal{L}_X[Y, Z] = [\mathcal{L}_X Y, Z] + [Y, \mathcal{L}_X, Z]$,
2. $\mathcal{L}_{[X,Y]} = [\mathcal{L}_X, \mathcal{L}_Y]$

From this last result, we can conclude that the bracket takes the form

$$\mathcal{L}_{[X,Y]} = \mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X. \tag{5.2}$$

We also conclude that if $[X, Y] = 0$ then $[\mathcal{L}_X, \mathcal{L}_Y] = 0$.

$$\begin{aligned} [\mathcal{L}_X, \mathcal{L}_Y](Z) &= \mathcal{L}_X\mathcal{L}_Y(Z) - \mathcal{L}_Y\mathcal{L}_X(Z) \\ &= [X, [Y, Z]] - [Y, [X, Z]] \\ &= [[X, Y], Z] = \mathcal{L}_{[X,Y]}(Z). \end{aligned}$$

For more specifics on the Lie derivative, see [LEE] page 103.

5.2 Metrics

Recall an inner product space is a vector space V equipped with a map $g : V \times V \to \mathbb{R}$ where if $v, w \in V$ then $g(\cdot, \cdot)$ is

1. $g(\cdot, \cdot) : V \otimes V \to \mathbb{R}$ is bilinear,
2. $g(v, w) = g(w, v)$ (symmetric),
3. $g(v, w) = 0$ for all w if and only if $v = 0$ (non-degenerate).

Most linear algebra texts require "$g(v, v) \geq 0$ for all v and $g(v, v) = 0$ if and only if $v = 0$ (positive definite)" instead of non-degenerate. We allow the (weaker) non-degenerate condition when building semi-Riemannian manifolds and require the (stronger) positive definite condition when building Riemannian manifolds. Note that positive definite implies non-degeneracy.

Notice that these hold true for the special case of the standard Euclidean inner product on $\mathbb{R}^n$, known as the dot product. We relax the requirement of positive definiteness to non-degeneracy (from Riemannian to semi-Riemannian) so as not to exclude metrics that occur in relativity.

Theorem 5.4. *There exists an orthonormal vector space V with a basis $\beta = \{e_1, ..., e_n\}$, such that*

$$g(e_i, e_j) = \epsilon_i \delta_{ij}$$

where $\epsilon_i = \pm 1$ and the number of -1's is called the signature or index.

Definition 5.5. *We say a vector $v \in V$ is spacelike if $g(v, v) > 0$ or $v = 0$. If $g(v, v) = 0$ and $v \neq 0$, that is v is a null vector, then v is light like. If $g(v, v) < 0$ then v is time like.*

Recall the definition of a metric from before. We will now extend this definition and define a semi-Riemannian metric g as a $(0, 2)$-tensor field on $\mathcal{M}$ (so $g_p : T_p\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R}$) such that at each point g is a symmetric non-degenerate bilinear form. Moreover, we require g has constant index as the point varies. In coordinates we write

$$g = g_{ij} dx^i \otimes dx^j. \tag{5.3}$$

That is for $v, w \in T_p\mathcal{M}$, the metric tensor g evaluates as

$$\begin{aligned} g(v, w) &= g(v^i \tfrac{\partial}{\partial x^i}, w^j \tfrac{\partial}{\partial x^j}) \\ &= g_{k\ell} dx^k (v^i \tfrac{\partial}{\partial x^i}) dx^\ell (w^j \tfrac{\partial}{\partial x^j}) \\ &= g_{k\ell} v^i w^j \delta_i^k \delta_j^\ell \\ &= g_{ij} v^i w^j \end{aligned}$$

Note that from the evaluation above that, $g_{ij} = g_{ji}$ and $\det(g_{ij}) \neq 0$. From our previous discussion on tensors, we know that g is an isomorphism that sends

$$g : V \to V^* \text{ by } v \mapsto g(v, \cdot) \tag{5.4}$$

and hence

$$v \in T_p\mathcal{M} \mapsto g(v,\cdot) \in T_p^*\mathcal{M} \tag{5.5}$$

$$v = v^i \frac{\partial}{\partial x^i} \mapsto g_{ij}v^j dx^i \tag{5.6}$$

The inverse of g is denoted as $[g_{ij}]^{-1} = g^{ij}$ where

$$g^{kl}g_{lj} = \delta^k_j. \tag{5.7}$$

Since the metric tensor connects a vector space with its dual in an isomorphic manner, it is also a tool for ”lowering” indicies. Consider $X^j\partial_j \in T_p\mathcal{M}$, then

$$X^j\partial_j = X \quad \text{implies} \quad g(X,\cdot) = g_{ij}dx^i(X) \otimes dx^j(\cdot) = g_{ij}x^i dx^j. \tag{5.8}$$

Therefore we can define the covariant components of X as

$$X_j = g_{ij}X^i \tag{5.9}$$

and hence we have

$$g^{ij}X_j = g^{ij}g_{ik}X^k = \delta^i_k X^k = X^i. \tag{5.10}$$

A technique for raising indicies is defined in a similar fashion.

5.3 Connections

In order to generalize our notion of a directional derivative from calculus, we require a useful technique for differentiating along generalized tangential structures, namely sections of a bundle. For example, differentiating along a section of the tangent bundle would give a directional derivative along a vector field. Connections do just this and differentiate sections in such a way that allows for the ”transporting” of a pair of vectors from one section to another.

Definition 5.6. *A connection D on a smooth manifold $\mathcal{M}$ is defined as a mapping*

$$D : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$$

Such that

1. *$D_V W$ is function-linear, or $\Omega^0(\mathcal{M})$-linear, in V,*
2. *$D_V W$ is $\mathbb{R}$-linear in W,*
3. *$D_V(fW) = (Vf)W + f(D_V W)$ for $f \in \Omega^0(\mathcal{M})$*

$D_V W$ (also denoted $\nabla_V W$) is known as the covarient derivative of W with respect to V for the connection D. From the first axiom, it follows that for $f_i \in \Omega^0(\mathcal{M})$

$$D_{f_1 V_1 + f_2 V_2}(W) = f_1 D_{V_1}(W) + f_2 D_{V_2}(W). \tag{5.11}$$

Recall that for a vector field $V = V^i \partial_i$, V acts on smooth functions as $V(f) = V^i \frac{\partial f}{\partial x^i}$. A choice of a connection is simply a chosen structure that induces a way of "pulling" curves from the base manifold to the total space parallel to sections at each point along the curve.

The notion of a connection is not restricted to the domain of vector fields in fact, we can extend our definition to tensor fields. Given a map ∇ : Tensor Fields $\to$ Tensor Fields, such that the type of the tensor field is preserved, we want to require the following properties to be satisfied:

1. ∇ is linear
2. $\nabla(A \otimes B) = \nabla(A) \otimes B + A \otimes \nabla(B)$
3. $\nabla(Contraction(A)) = Contraction(\nabla(A))$

If all three properties are satisfied then ∇ is a tensor derivation. We denote the component of ∇ acting on (r, s) tensors by ∇^r_s.

Example 5.7. *Consider $\nabla^0_0 : \Omega^0(\mathcal{M}) \to \Omega^0(\mathcal{M})$.*

1. *∇^0_0 is linear since since we can pull out scalars and distrubute over smooth functions,*

2. $\boldsymbol{\nabla}_0^0(fg) = g\boldsymbol{\nabla}_0^0(f) + f\boldsymbol{\nabla}_0^0(g)$ *which tells us* $\boldsymbol{\nabla}_0^0 \in \mathrm{Der}(\Omega^0(\mathcal{M})) = \mathfrak{X}(\mathcal{M})$.

From this we conclude that there exists a vector field X such that $\boldsymbol{\nabla}_0^0(f) = X(f)$.

Now consider a general (r, s) tensor $A = A(\theta^1, ...\theta^r, X_1, ..., X_s)$. Then

$$\boldsymbol{\nabla}(A(\theta^1, ...\theta^r, X_1, ..., X_s)) =$$
$$(\boldsymbol{\nabla} A)(\theta^1, ...\theta^r, X_1, ..., X_s) + \sum A(..., \boldsymbol{\nabla}(\theta^i), ...) + \sum A(..., \boldsymbol{\nabla} X_j, ...). \tag{5.12}$$

To sketch a proof of this, we will consider the case for $r = s = 1$ for simplicity. To begin we use our third requirement for a tensor derivation and show

$$\begin{aligned}
\boldsymbol{\nabla}(A(\theta, X)) &= \\
\boldsymbol{\nabla}(contraction(A \otimes \theta \otimes X)) = contraction(\boldsymbol{\nabla}(A \otimes \theta \otimes X)) &= \\
contraction(\boldsymbol{\nabla}(A) \otimes \theta \otimes X + A \otimes \boldsymbol{\nabla}(\theta) \otimes X + A \otimes \theta \otimes \boldsymbol{\nabla}(X)) &= \\
\boldsymbol{\nabla}(A)(\theta, X) + A(\boldsymbol{\nabla}(\theta), X) + A(\theta \otimes \boldsymbol{\nabla}(X))
\end{aligned}$$

Notice that we have $\boldsymbol{\nabla}_0^0(A(\theta, X)) = \boldsymbol{\nabla}_1^1(A)(\theta, X) + A(\boldsymbol{\nabla}_0^1(\theta), X) + A(\theta \otimes \boldsymbol{\nabla}_1^0(X))$. From which we conclude that $\boldsymbol{\nabla}_1^1$ is determined by $\boldsymbol{\nabla}_0^0$, $\boldsymbol{\nabla}_1^0$, and $\boldsymbol{\nabla}_0^1$. Likewise, the same is true for $\boldsymbol{\nabla}_s^r$. Furthermore,

$$(\boldsymbol{\nabla}_0^1\theta)(X) = \boldsymbol{\nabla}_0^0(\theta(X)) - \theta(\boldsymbol{\nabla}_1^0(X))$$

So if we know $\boldsymbol{\nabla}$ is a tensor derivation and we know $\boldsymbol{\nabla}_0^0$ and $\boldsymbol{\nabla}_1^0$, all other components follow. We only need to know how $\boldsymbol{\nabla}$ acts on functions and vector fields.

Definition 5.8. *The torsion tensor τ is defined as the mapping*

$$\tau : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$$

such that for a connection $\boldsymbol{\nabla}$ and vector fields X and Y

$$\tau(X, Y) = \boldsymbol{\nabla}_X Y - \boldsymbol{\nabla}_Y X - [X, Y]$$

Notice if our connection is torsion free, that is $\tau(X, Y) = 0$, then the Lie bracket is defined as

$$[X, Y] = \boldsymbol{\nabla}_X Y - \boldsymbol{\nabla}_Y X. \tag{5.13}$$

Definition 5.9. *We define metric compatibility with the chosen connection, vector fields X, Y, Z, and the metric g as*

$$X(g(Y,Z) = g(\nabla_X(Y), Z) + g(Y, \nabla_X(Z))$$

Such a connection ∇ is called a metric connection.

Example 5.10. *The Levi-Civita Connection: For a given semi-Riemannian manifold $\mathcal{M}$, there exists a unique torsion free, metric connection ∇. We call this connection the Levi-Civita Connection.*
The Levi-Civita connection ∇ satisfies the Koszul formula, whose validity is proven in [ON].
Let X, Y, Z be arbitrary vector fields on $\mathcal{M}$ with a metric g, then

$$2g(\nabla_X Y, Z) = Xg(Y,Z) + Yg(Z,X) - Zg(X,Y) + g(Z,[X,Y]) - g(Y,[X,Z]) - g(X,[Y,Z]).$$

The Koszul formula must be satisfied for any connection on a semi-Riemannian manifold. Note from metric compatibility that we must have

$$X(g(Y,Z)) = g(\nabla_X(Y), Z) + g(Y, \nabla_X(Z))$$

$$Y(g(X,Z)) = g(\nabla_Y(X), Z) + g(X, \nabla_Y(Z))$$

$$Z(g(X,Y)) = g(\nabla_Z(X), Y) + g(X, \nabla_Z(Y))$$

We then find from combining these three expressions

$$X(g(Y,Z)) + Y(g(X,Z)) - Z(g(X,Y)) =$$

$$g(\nabla_X(Y), Z)) + g(Y, \nabla_X(Z)) + g(\nabla_Y(X), Z) + g(X, \nabla_Y(Z))$$

$$-g(\nabla_Z(X), Y) - g(X, \nabla_Z(Y)).$$

By using the linearity and symmetry of the metric, the above sum becomes

$$X(g(Y,Z)) + Y(g(X,Z)) - Z(g(X,Y)) =$$
$$g(Z, \nabla_Y(X) - \nabla_X(Y)) + g(Y, \nabla_X(Z) - \nabla_Z(X)) +$$
$$g(X, \nabla_Y(Z) - \nabla_Z(Y) + 2g(\nabla_X Y, Z)$$

After applying our condition for zero torsion to swap out the Lie bracket and rearranging we arrive at the familiar

$$2g(\nabla_X Y, Z) = Xg(Y,Z) + Yg(Z,X) - Zg(X,Y) + g(Z,[X,Y]) - g(Y,[X,Z]) - g(X,[Y,Z]).$$

Since we need only to know $\nabla_X Y$ *to calculate* $g(\nabla_X Y, Z)$ *for any* Z *and vice-versa, we can conclude that if* ∇ *exists then it must be unique. The only appearance of* $\nabla_X Y$ *is on the left hand side. So the connection inner producted with* Z *can be computed in terms of the metric and Lie derivative. But once we know* $g(\nabla_X Y, Z)$ *for all* Z *then we know* $\nabla_X Y$*. Specifically, g(A,Z)=g(B,Z) for all* Z *implies* $g(A-B,Z) = 0$ *for all* Z *implies* $A - B = 0$ *(by non-degeneracy) so therefore* $A = B$*.*

The Levi-Civita connection is also used to define derivatives along paths or curves, giving rise to the pullback connection and parallel transport.

Connections have many uses in mathematics and physics. In fact, the standard model for particle physics is founded on the idea of a gauge field, which can be viewed as sections of the associated bundle. Specifically, a gauge field theory is defined to be the vector bundle $(E, \mathcal{M}, V, \pi)$ over a smooth manifold $\mathcal{M}$ for which the vector space V carries a representation of a Lie group G, known as the gauge group. Sections of this bundle are called matter fields and connections on this bundle are called gauge potentials, which are local connection forms. In this theory, the nucleon field is also a section of a vector bundle over a four dimensional Lorentzian manifold whose fibers are two dimensional vector spaces over $\mathbb{C}$. It turns out that since each complex vector space carries equivalent representations of $SU(2)$ there are invariances under local group transformations therefore a covarient derivative is required. The notions of representation theory, group transformations, and Lie groups are unfortunately just beyond the scope of this paper. For a brief outline of gauge theory and how the mathematics in this paper applies, see [R] exercise 7.10, page 190.

5.4 Covariant Derivative

There is generally much confusion surrounding the difference between the Levi-Civita connection and the covariant derivative. However the notions of connections and covariant differentials are intimately intertwined, so much that they are often used synonymously. To recap, a choice of a connection is a specific choice of a mapping, as outlined in the previous section. Given said choice of a connection, we may define a covariant derivative. Here we will first view the covariant derivative from the bundle perspective.

Definition 5.11. *Let E be a vector bundle over a manifold $\mathcal{M}$. An affine connection or covariant derivative on E is a map*

$$\boldsymbol{\nabla} : \mathfrak{X}(\mathcal{M}) \times \Gamma(E) \to \Gamma(E)$$

where $\boldsymbol{\nabla}(X, \mathfrak{s})$ is denoted as $\boldsymbol{\nabla}_X \mathfrak{s}$ and satisfies the following properties

1. $\boldsymbol{\nabla}_f \mathfrak{s} = f \boldsymbol{\nabla}_X \mathfrak{s}$ *for all* $f \in \Omega^0(\mathcal{M}),\ X \in \mathfrak{X}(\mathcal{M})$, *and* $\mathfrak{s} \in \Gamma(E)$ *;*
2. $\boldsymbol{\nabla}_{X_1+X_2} \mathfrak{s} = \boldsymbol{\nabla}_{X_1} \mathfrak{s} + \boldsymbol{\nabla}_{X_2} \mathfrak{s}$*;*
3. $\boldsymbol{\nabla}_X(\mathfrak{s}_1 + \mathfrak{s}_2) = \boldsymbol{\nabla}_X(\mathfrak{s}_1) + \boldsymbol{\nabla}_X(\mathfrak{s}_2)$
4. $\boldsymbol{\nabla}_X(f\mathfrak{s}) = (Xf)\mathfrak{s} + f\boldsymbol{\nabla}_X \mathfrak{s}$

Given a fixed $X \in \mathfrak{X}(\mathcal{M})$, the mapping

$$\boldsymbol{\nabla}_X : \Gamma(E) \to \Gamma(E) \tag{5.14}$$

is called the covariant derivative with respect to X. Note that the covariant derivative can also be defined as the map $\boldsymbol{\nabla} : T\mathcal{M} \times \Gamma(E) \to \Gamma(E)$ where for some $v \in T_p\mathcal{M}$, $\boldsymbol{\nabla}_v \mathfrak{s} = (\boldsymbol{\nabla}_X \mathfrak{s})(p)$ given any extension of $v \in T_p\mathcal{M}$ to X such that $X\Big|_p = v$.

Definition 5.12. *Let $x^1, ..., x^n$ be the standard coordinates of $\mathbb{R}^n$, then if V and W are vector fields on $\mathbb{R}^n$, expressed as $V = V^i\partial_i$ and $W = W^i\partial_i$ respectively, then*

$$\boldsymbol{\nabla}_V W = V(W^i)\partial_i$$

is a vector field called the natural covarient derivative of W with respect to V.

Definition 5.13. *Let $\boldsymbol{\nabla}$ be the covariant derivative and $x^1, ..., x^n$ be local coordinates. We define*

$$\boldsymbol{\nabla}_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma^k_{ij} \frac{\partial}{\partial x^k}$$

where the Γ^k_{ij} are smooth functions and called the Christoffel symbols (of the second kind) or connection coefficients of the connection $\boldsymbol{\nabla}$. Note that these coefficients are sometimes denoted as $\{{k \atop ij}\}$.

Consider the Lie derivative and recall that mixed partials commute

$$0 = \left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = \boldsymbol{\nabla}_{\frac{\partial}{\partial x^i}} \left(\frac{\partial}{\partial x^j}\right) - \boldsymbol{\nabla}_{\frac{\partial}{\partial x^j}} \left(\frac{\partial}{\partial x^i}\right) = \left(\Gamma^k_{ij} - \Gamma^k_{ji}\right) \frac{\partial}{\partial x^k} \tag{5.15}$$

So $\Gamma^k_{ij} - \Gamma^k_{ji}$ is the k-th component of the zero vector field. Therefore,

$$\Gamma^k_{ij} = \Gamma^k_{ji} \tag{5.16}$$

Recall that we may use the metric tensor to lower indices, such that (thus switching from Christoffel symbols of the second to the first kind):

$$\Gamma^\ell_{ij} g_{\ell k} = \Gamma_{jik}. \tag{5.17}$$

We now return to the Koszul equation and substitute $\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}$ for X, Y, Z respectively to obtain

$$2g(\boldsymbol{\nabla}_{\frac{\partial}{\partial x^i}} \tfrac{\partial}{\partial x^j}, \tfrac{\partial}{\partial x^k}) = \tfrac{\partial}{\partial x^i} g(\tfrac{\partial}{\partial x^j}, \tfrac{\partial}{\partial x^k}) + \tfrac{\partial}{\partial x^j} g(\tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^k}) - \tfrac{\partial}{\partial x^k} g(\tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^j})$$

$$+ g(\tfrac{\partial}{\partial x^k}, [\tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^j}]) - g(\tfrac{\partial}{\partial x^j}, [\tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^k}]) - g(\tfrac{\partial}{\partial x^i}, [\tfrac{\partial}{\partial x^j}, \tfrac{\partial}{\partial x^k}]).$$

$$2g(\Gamma^{\ell}_{ij}\partial_{\ell}, \partial_k) = \partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij} + 0 - 0 - 0$$

$$2\Gamma^{\ell}_{ij} g_{\ell k} = \partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}$$

$$2\Gamma^{m}_{ij} = 2\Gamma^{\ell}_{ij}\delta^{m}_{\ell} = 2\Gamma^{\ell}_{ij} g_{\ell k} g^{km} = (\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}) g^{km}$$

Simplifying further we obtain a nice way of calculating the Christoffel symbols (of the first kind):

$$\Gamma_{ijk} = \frac{1}{2}\left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\right) \tag{5.18}$$

Example 5.14. *Cylindrical example. Consider $\mathbb{R}^3$ with coordinates (x, y, z) and let the basis $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\}$ be denoted as $\{\partial_x, \partial_y, \partial_z\}$. We define the metric as $ds^2 = g(\partial_i, \partial_j)dx^i dx^j = dx^2 + dy^2 + dz^2$ (note here the superscript 2 denotes exponentiation). Then transforming to standard cylindrical coordinates such that $x = r\cos(\theta)y = r\sin(\theta)$ and $z = z$, we find*

$$\tfrac{\partial}{\partial r} = \cos(\theta)\tfrac{\partial}{\partial x} + \sin(\theta)\tfrac{\partial}{\partial y}\,,\;\; \tfrac{\partial}{\partial \theta} = -r\sin(\theta)\tfrac{\partial}{\partial x} + r\cos(\theta)\tfrac{\partial}{\partial y}\,,\;\; \tfrac{\partial}{\partial z} = \tfrac{\partial}{\partial z}.$$

We then find the components of the metric to be

$$g(\partial_r, \partial_r) = \cos^2(\theta) + \sin^2(\theta) = 1$$

$$g(\partial_r, \partial_\theta) = 0$$

$$g(\partial_\theta, \partial_\theta) = r^2$$

$$g(\partial_z, \partial_z) = 1$$

$$\textit{otherwise, } g(\partial_i, \partial_j) = 0.$$

Now define $y^1 = r, y^2 = \theta, y^3 = z$, then the components of the metric tensor are simply $g_{11} = g_{33} = 1$,

$g_{22} = r^2$, and $g_{ij} = 0$ otherwise. We then can calculate the Christoffel symbols as

$$\Gamma^1_{22} = \frac{1}{2} g^{1m} \left(\frac{\partial g_{2m}}{\partial y^2} + \frac{\partial g_{2m}}{\partial y^2} - \frac{\partial g_{22}}{\partial y^m} \right).$$

Note that this is non-zero if and only if $m = 1$, therefore

$$\Gamma^1_{22} = \frac{1}{2} g^{11} \left(\frac{\partial g_{21}}{\partial y^2} + \frac{\partial g_{21}}{\partial y^2} - \frac{\partial g_{22}}{\partial y^1} \right) = -r.$$

Likewise

$$\Gamma^2_{21} = \Gamma^2_{12} = \frac{1}{r}$$

$$\textit{otherwise,} \quad \Gamma^m_{ij} = 0.$$

Since the connection coefficients have been obtained, we can now define the covariant derivative in terms of the chosen coordinates

$$\boldsymbol{\nabla}_{\partial_\theta}(\partial_\theta) = \Gamma^k_{22} \partial_k = -r \partial_r$$

$$\boldsymbol{\nabla}_{\partial_\theta}(\partial_r) = \Gamma^k_{21} \partial_k = \frac{1}{r} \partial_\theta$$

At last we have completed our first geometric computation. By this point it is fair to note that we have hardly even begun an expedition into differential geometry. In fact, it would be wrong to say that we have even begun to scratch the surface. The proper next step in exploring this topic would be to delve into the Riemann curvature tensor and begin deciphering the notion of parallelism. This would leave the reader at a place where they could likely pick up a book on general relativity or another area of interest and begin without much trouble. There is still much left to be explored.

References

[EW] K. Erdmann and M. J. Wildon. *Introduction to Lie Algebras*. Springer, London, 2006.

[F] Theodore Frankel. *The Geometry of Physics, An Introduction*. Cambridge University Press, Cambridge, 1997.

[HK] K. Hoffman and R. Kunze. *Linear Algebra*. Prentice-Hall, New Jersey, 1971.

[K] Erwin Kreyszig. *Introductory Functional Analysis with Applications*. John Wiley & Sons. Inc., United States of America, 1978.

[LEE] Jeffrey M. Lee. *Manifolds and Differential Geometry*. American Mathematical Society, Providence, RI, 2009.

[M] James R. Munkres. *Topology*. Pearson, New York, 2018.

[NLab] nLab. Differentiable Manifold Image. (2016, August 14) Retrieved April 23, 2019, from https://mathtuition88.com/2016/08/14/differentiable-manifold/

[ON] Barrett O'Neill. *Semi-Riemannian Geometry With Applications to Relativity*. Academic Press, San Diego, 1983.

[R] Paul Renteln. *Manifolds, Tensors, and Forms*. Cambridge University Press, Cambridge, 2014.

[Sk] A. Skopenkov. "Embedding and knotting of manifolds in Euclidean spaces" in *Surveys in Contemporary Mathematics*. Ed. N. Young and Y. Choi, London Math. Soc. Lect. Notes. 347 (2): 248–342, 2008.

[Sp] Michael Spivak. *A Comprehensive Introduction to Differential Geometry* (Vol. 1). Publish or Perish, Inc, Houston, 2005.