Tensors and Manifold Theory

By
Aeron McConnell

Departmental 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 in General Mathematics
May, 2021

Approved by:

Dr. William Cook, Thesis Director

Dr. Quinn Morris, Second Reader

Dr. Eric Marland, Chair

Acknowledgments

I want to take a moment here to thank my thesis advisor Dr. Bill Cook. I greatly appreciate the countless hours he spent teaching me and answering my interminable stream of questions. Dr. Cook made writing a thesis fun when it easily could have been a chore.

I also want to thank Dr. Quinn Morris for being a great second reader and for being willing to look over way more of my thesis materials than he should have had to look over.

Thanks for everything!

Contents

1 Introduction

The study of geometry has been around for as long as people have thought about shapes, but it wasn't until the 17-1800's that people like Gauss, Bolyai and Lobachevsky started seriously studying non-Euclidean geometries. It was in the research of this era that the first ideas of manifolds started to take shape [W]. Gauss was also the first person to discuss some of the big theorems we usually see at the end of a standard multi-variable calculus class. Specifically, in 1813 he proposed a theorem which is a version of our modern divergence theorem or Gauss' theorem. He continued publishing other special cases of that theorem over the next few decades despite the fact that Michael Ostrogradsky had proven the general theorem in 1826 [MAA]. The concept of manifolds started gaining more attention in the next few decades as Niels Abel and Carl Jacobi considered certain complex manifolds and Bernhard Riemann worked tirelessly to generalize surfaces to higher dimensions in what he called Mannigfaltigkeit, the origin of our term "manifold". In 1895 Poincarè published a seminal paper titled "Analysis Situs" in which he defines a differentiable manifold and around the same time (1890-1892) Gregorio Ricci-Curbastro developed tensor calculus. These fields continued growing in the early 20th century with developments such as Einstein's use of tensors in General Relativity and the discovery of more exciting theorems such as the Whiteney Embedding Theorem for manifolds. At this point the study of manifolds and tensors had become so far-reaching and rich that it is not possible to easily summarize: Manifolds became fundamental to areas of physics, applied and pure math, engineering, and even fields like chemistry and biology. Due to the use of manifolds and tensors throughout much of modern science, especially in areas where complicated surfaces and objects arise, some understanding of these concepts would be beneficial to most people in STEM and as such we present a summary of some of the basics of tensors and manifold theory in this paper along with some brief discussion of the Generalized Stokes Theorem which is arguably one of the biggest theorems in calculus.

2 Review of necessary fundamentals

Before we can get to talking about manifolds and tensors it would be prudent to review some of the basics that underlie these topics. These will include defining vector spaces and dual spaces along with some terminology and notation that will be useful to us later. Much of the information in this section comes from Paul Renteln's "Manifolds, Tensors, and Forms" [R].

2.1 Vector spaces

A basic understanding of vectors is expected at this level, but due to the importance of vector spaces and basis sets in essentially everything after this point we should spend some time making sure we are familiar with them.

Definition 2.1.1: Vector Spaces

A **vector space** V over a field $\mathbb{F}$ is defined to be a nonempty set that is an Abelian group under addition and given $a, b \in \mathbb{F}$ and $v, w \in V$ we have:

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

Note that being an Abelian group under addition means that the set V abides by the following properties for all $v, w, x \in V$:

1. Closure,
   $v + w \in V$
2. Associativity,
   $(v + w) + x = v + (w + x)$
3. Identity,
   $\exists 0 \in V$ such that
   $0 + v = v + 0 = v$
4. Inverses,
   $\exists -v \in V$ such that
   $v + (-v) = (-v) + v = 0$
5. Commutativity,
   $v + w = w + v$

We will most commonly deal with vector spaces over the real numbers which means our vector spaces will look like $\mathbb{R}^n$ and contain n-tuples of real numbers, the vectors, along with individual real numbers, the scalars. Let us consider a less abstract case as an example.

Example 2.1.1

If we apply the third property listed in definition 2.1.1 to some randomly chosen elements of $V = \mathbb{R}^2$, say $a = 2$, $b = 3$, and $v = (3, 1)$, we see that the elements of vector spaces behave in fairly intuitive ways:

$$(a+b)v = (2+3)(3,1) = 5(3,1) = (15,5) = (6,2)+(9,3) = 2(3,1)+3(3,1) = av+bv$$

It would be nice to have a more systematic way to think about and understand vectors in vector space, but to do this we will first need a couple of definitions. First, since this terminology is relevant to the next definition, given vectors $v_1, \ldots, v_k$ in a vector space V and scalars $c_1, \ldots, c_k \in \mathbb{F}$, we says $c_1 v_1 + \cdots + c_k v_k$ is a finite linear combination of the vectors $v_1, \ldots, v_k$. Going forward when we say "linear combination" we mean finite linear combination.

Definition 2.1.2: Linear Independence

A set of vectors $S \subset V$ are said to be **linearly independent** if, for any finite subset of vectors $\{v_1, \ldots, v_k\} \subseteq S$ and set of scalars $\{a_1, \ldots, a_k\} \subset \mathbb{F}$, then:

$$\sum_{i=1}^{k} a_i v_i = 0 \qquad \text{implies that} \qquad a_i = 0 \text{ for all } i$$

Briefly, only a trivial linear combination of elements of S can yield 0.

It is easily shown that a set of vectors is linearly independent if no vector in that set can be written as a linear combination of the other vectors or in other words as the sum of the other vectors times scalars from $\mathbb{F}$.

Definition 2.1.3: Spanning Sets

A set of vectors $W \subset V$ is called a **spanning set** for V, or is said to **span** V, if every vector in V can be written as a linear combination of vectors from W. In fact, given $S \subseteq V$, we denote the set of linear combinations of vectors drawn from S by $\text{Span}(S)$. Here S spans $\text{Span}(S)$.

Definition 2.1.4: Basis Sets

A set, β, of linearly independent vectors that span V is called a **basis** for the vector space V.

It can be shown (Page 154 of [J]) that any two bases for a vector space V must have the same cardinality. Therefore, we call the cardinality of some basis (hence every

basis) for V the dimension of V and denote it by $\dim(V)$.
Now we have access to a more fundamental description of these vector spaces using the definition of a basis set. The elements of these basis sets are very useful because they act like the atoms in the universe of our vector space. As an example consider the real numbers again:

Example 2.1.2

The standard basis for $\mathbb{R}^3$ is the set $\{e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1)\}$. Since the basis contains three elements we say that the vector space has dimension 3. Now it should also be clear that any element of the $\mathbb{R}^3$ vector space can be constructed component-wise from these basis elements as: $v = v_1e_1 + v_2e_2 + v_3e_3 = (v_1, v_2, v_3)$ where $v_1, v_2, v_3 \in \mathbb{R}$. Finally, note that there is nothing special about $\mathbb{R}^3$ here and this could be generalized to any $\mathbb{R}^n$ containing n basis elements. This can also be generalized to fields other than the real numbers and bases other than the standard basis although these take a little more notation to avoid ambiguity.

2.2 Dual Spaces

Although it will not be clear for a little while why we are defining these dual vectors and dual spaces, they become very important when we want to understand tensors in more depth. For now all we need to know is that these odd types of objects exist and are defined as follows.

Definition 2.2.1: Dual Space

The set, V^*, of all linear maps $f : V \to \mathbb{F}$ is called the **dual space** of V. It can also be written as $\mathrm{Hom}(V, \mathbb{F})$ since it is the set of all linear homomorphisms from V into $\mathbb{R}$. Each of those maps $f \in V^*$ is called a linear functional and these are unsurprisingly linear objects: $f(av+w) = af(v)+f(w)$ for all $v, w \in V$ and $a \in \mathbb{F}$. Also, the set of such maps forms a vector space with its elements commonly called **dual vectors** or covectors.

When V is finite dimensional, one can show that V^{**} is canonically isomorphic to V via the evaluation map (i.e., $ev : V \to V^{**}$ where $ev(v) : V^* \to \mathbb{F}$ is defined by $ev(v)(f) = f(v)$ for any $f \in V^*$). Therefore, given $f \in V^*$ and $v \in V$, one may consider v as a double dual vector and thus $f(v)$ can be view as v plugged into f or as f plugged into v. To highlight this ambiguity we often write $\langle f, v\rangle$ for $f(v)$. This dual pairing notation puts vectors and dual vector (covectors) on more-or-less equal footing.
As the name of the elements of this dual space imply, there is a connection between vectors and covectors. Suppose V is a finite dimensional vector space with $\dim(V) = n$. Given a basis consisting of $\{e_i \,|\, i = 1, \ldots, n\}$ for V then we can define a collection of

dual vectors $\{\theta^j \mid j = 1, \ldots, n\}$ as follows:

$$\langle e_i, \theta^j \rangle = \langle \theta^j, e_i \rangle = e_i(\theta^j) = \delta_i^j$$

where δ_{ij} or δ_i^j is the Kronecker delta function and therefore equals 1 when $i = j$ and 0 otherwise. It can be shown that $\{\theta^j \mid j = 1, \ldots, n\}$ is a basis for V^*. We call this the basis dual to $\{e_i \mid i = 1, \ldots, n\}$, or briefly, a dual basis.
Given, $v \in V$ and $f \in V^*$ along with basis $\beta = \{e_1, \ldots, e_n\}$ for V and dual basis $\beta^* = \{\theta^1, \ldots, \theta^n\}$ for V^*, we have v is a linear combination of the elements of β, $v = v^1 e_1 + \cdots + v^n e_n$ and f is a linear combination of elements of β^*, $f = f_1 \theta^1 + \cdots + f_n \theta^n$. If we have an understood basis, we will let v^i denote these coordinates for v and f_j denote these coordinates for f. But what are coordinates? Well it turns out that there has been some shady notation going on and we are just ignoring some of the inconsistencies. Let us now more rigorously define what we mean by coordinates and why we are writing things in the way we do. Consider a point in $\mathbb{R}^n$. Usually we would denote this as an n-tuple like $(x^1, x^2, \cdots, x^n)$ where each of the x^i's is a **coordinate** of that point. But, in other contexts we use these notations to denote **coordinate functions**: $x^i : \mathbb{R}^n \to \mathbb{R}$ defined as $x^i(p^1, \cdots, p^n) = p^i$ or $x^i(p) = p^i$. Juxtaposing these notations we get something bizarre like $x^i(x) = x^i$ where x^i is a coordinate function on the left and a coordinate/point on the right. We are addressing this now not for the purpose of providing some beautiful notation that solves these problems, but instead to say essentially "sorry, both notations are common, deal with it". Generally it should be clear from context what is meant by such a coordinate function/coordinate expression, but we will try to avoid it nonetheless.
To those unfamiliar with Einstein notation and the differences between covariant/contravariant components the previous section has been rather strange. To address this we need to understand these notations better.

Einstein Notation and Covariant/Contravariant Components

The above discussion uses superscripts commonly reserved for exponents and this is not very intuitive upon first sight, but rest assured it is for a good reason. If we consider changing the basis for a vector we find that the basis elements of that vector transform like basis covectors, that is, contravariantly or "against the basis" of the vector space. Conversely, if we preform a change of basis for a covector we see that the components of the covector transform like the vector space basis elements, that is, covariantly or "with the basis" of the vector space. As such this paper will use lower indices to denote contravariant components for vectors like e_i and upper indices to denote covariant components for covectors like θ^j. While this is useful in understanding vectors and covectors under a change of basis, the true power of this notation is how it allows us to interpret expressions. Not only does this notation make it much easier to identify whether something is in the vector space or the dual space, it also pairs very nicely with Einstein notation. Due to the vast number of complicated summations associated with tensors Einstein notation is commonly used to greatly simplify equations. The main idea of Einstein notation is that when the same index shows up more than once in any term of a sum, the sum is implied so the symbol may be omitted. An example of this is, imagine we want to compute $f(v)$ where $f \in V^*$ and $v \in V$. This expression can then be written as we usually see it using a summation:

$$f(v) = \sum_{j=1}^{n} \sum_{i=1}^{n} f_j v^i$$

Or it could be written in Einstein notation using $f = f_j \theta^j$, $v = v^i e_i$, and $\theta^j(e_i) = \delta_i^j$:

$$f(v) = f_j \theta^j (v^i e_i) = f_j v^i \theta^j (e_i) = f_j v^i \delta_i^j = f_i v^i$$

3 Tensors

Tensors are fundamentally multilinear objects meaning that they are linear in multiple "slots". They also have many possible representations giving them some serious versatility in uses including as tensor products, multilinear maps, and arrays among other representations. Their wide use throughout much of physics, especially in discussions about relativity and manifolds, makes them thoroughly worth studying.

3.1 Basic Definitions

In the basic description of tensors they are essentially an expanded and "higher dimensional" version of vectors although they don't have to be higher dimensional since technically vectors are "1-tensors" and scalars are "0-tensors". To better understand the basics of tensors here is a definite example of working with tensors and their different forms.

Example 3.1.1

Imagine we want to deal with tensors of the form $v \otimes w \in \mathbb{R}^2 \otimes \mathbb{R}^3$. If we assume that the canonical bases of $\mathbb{R}^2$ and $\mathbb{R}^3$ are $\{e_1, e_2\}$ and $\{e_1, e_2, e_3\}$ respectively then the basis for this vector space, $\mathbb{R}^2 \otimes \mathbb{R}^3$, is

$$\beta = \{e_1 \otimes e_1,\ e_1 \otimes e_2,\ e_1 \otimes e_3,\ e_2 \otimes e_2,\ e_2 \otimes e_3,\ e_3 \otimes e_3\}$$

which includes all combinations of the original basis elements of $\mathbb{R}^2$ and $\mathbb{R}^3$. Taking an arbitrary element from this space, called a 2-tensor or tensor of order two because it is the tensor product of two vectors, we can demonstrate how it exhibits multilinearity by showing the linear property it has in its first vector:

$$(3,2) \otimes (1,2,3) = 3(1,0) \otimes (1,2,3) + 2(0,1) \otimes (1,2,3)$$

Similarly these pieces can be expanded in the second vector slot and this actually shows what the tensor looks like in terms of its basis elements:

$$(3,2) \otimes (1,2,3) = 3e_1 \otimes e_1 + 6e_1 \otimes e_2 + 9e_1 \otimes e_3 + 2e_2 \otimes e_1 + 4e_2 \otimes e_2 + 6e_2 \otimes e_3$$

From here is should be fairly apparent how tensors like this can be expressed as indexed arrays similar to matrices:

$$(3,2) \otimes (1,2,3) = \begin{bmatrix} 3 & 6 & 9 \\ 2 & 4 & 6 \end{bmatrix}$$

In general the multilinear property looks like:

$$T(v_1, ..., av + w, ..., v_n) = aT(v_1, ..., v, ..., v_n) + T(v_1, ..., w, ..., v_n)$$

Where T is a multilinear map from vector spaces $V_1, \cdots, V_n$ to the vector space W, a is a constant, and $v, w \in V_i$.

While this is a useful example it is not a substitute for definitions and properties so listed below are a few properties of tensors. Given tensors R with order r, S and T with order s, and constants α and β we can write:

- The set of all tensors $\mathcal{R} = \cup_r R^r$ forms an Algebra
- The tensor $R \otimes S$ has order $r + s$
- The expression $\alpha S + \beta T$ is a tensor of order s
- Tensors are associative but not commutative
- Some rearranging and distributive laws hold:
    1. $T \otimes (\alpha S) = (\alpha T) \otimes S = \alpha(T \otimes S)$
    2. $R \otimes (S + T) = R \otimes S + R \otimes T$
    3. $(S + T) \otimes R = S \otimes R + T \otimes R$

The main downside to a representation like Example 2.1.1 is that it obscures how tensors are basis independent. Another way we can think about tensors is as a multilinear map satisfying a universal property:

Definition 3.1.1: Tensors

Given vector spaces $V_1, V_2, \cdots, V_n$ and a multilinear map $T : V_1 \times V_2 \times \cdots \times V_n \to W$ mapping to a vector space W then there exists a unique linear map $\hat{T} : V_1 \otimes V_2 \otimes \cdots \otimes V_n \to W$ such that $T(V_1, V_2, \cdots, V_n) = \hat{T}(V_1 \otimes V_2 \otimes \cdots \otimes V_n)$. Multilinear mapping objects of this form, $V_1 \otimes V_2 \otimes \cdots \otimes V_n$, are called **tensors**. Consider this statement represented in the diagram below:

$$\begin{tikzcd}
V_1 \otimes V_2 \otimes \cdots \otimes V_n \arrow[dr, dashed, "\hat{T}"] & \\
V_1 \times V_2 \times \cdots \times V_n \arrow[u, "\otimes"] \arrow[r, "T"'] & W
\end{tikzcd}$$

Multilinear $\longrightarrow$ Linear $\dashrightarrow$

3.2 General Tensors

While the above discussions of tensors are useful in understanding tensors to some extent, modifications need to be made in order to also include dual objects.

Definition 3.2.1: General Tensors

A **general tensor** consists of contravariant and covariant pieces combined through a tensor product. Take a vector space V with basis $\{e_i\}$ and its dual space V^* with basis $\{\theta^j\}$ then a general (r,s)-tensor will have the form:

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

Note that this is using Einstein summation notation and has excluded the summation symbol. These tensors belong to a tensor product space that looks like:

$$\mathcal{T}^r_s = \underbrace{V \otimes V \otimes \cdots \otimes V}_{r \text{ times}} \otimes \underbrace{V^* \otimes V^* \otimes \cdots \otimes V^*}_{s \text{ times}} = V^{\otimes r} \otimes (V^*)^{\otimes s}$$

These tensors actually give us a wide range of possible representations depending on how we want to represent them. The following example will show how tensors can be thought of as many different maps.

Example 3.2.1

Consider the mapping $T : V \otimes V \otimes V^* \to \mathbb{R}$. This is a (2,1)-tensor with elements of the form $v \otimes w \otimes f$ but we could instead leave one or more "slots" open and change how this tensor maps objects. If we plug in $v \otimes w$, leaving open the dual vector slot, we would instead get the mapping $T(v, w, \cdot)$ which takes in a dual vector and kicks out a scalar, so $T : V \otimes V \to V^{**}$. Identifying V^{**} and V makes this the map $T : V \otimes V \to V$. In coordinates, $T^k_{ij} v^i w^j$ looks like a vector (note the hanging upper index k). But this is not the only possibility, in fact there are quite a few interpretations for our tensor T:

$$T : V \otimes V \otimes V^* \to \mathbb{R} \qquad T : V \otimes V \to V \qquad T : V \otimes V^* \to V^*$$

$$T : V \to V \otimes V^* \qquad T : V^* \to V^* \otimes V^* \qquad T \in V^* \otimes V^* \otimes V$$

For greater detail on "mixed" tensors see T. Frankel's "The Geometry of Physics" section 2.4c [F].

4 Manifolds

In order to extend the use of powerful calculus tools to more complicated objects we need to discuss manifolds. Manifolds allow us to consider complicated objects made from pieces of $\mathbb{R}^n$ by giving us a way to break them down and understand the simpler pieces involved. Yet manifolds are inherently topological and we need to have some basis in topology in order to full fully explore manifolds.

4.1 Topology

Topology takes sets of points and gives them some form of geometrically inspired "rules" allowing us to study properties that are inherent to the set. A topology on a set tells us which subsets of this set are open, or belong in the topology. Much of the information in this section comes from chapter 2 of [M].

Definition 4.1.1: Topology

A **topology** $\mathcal{T}$ on a set X consists of all subsets of X which are open. Open sets satisfy the following three conditions:

1. The full set, X, and the empty set, $\emptyset$, are both open sets
2. Arbitrary unions of open sets are open
3. Finite intersections of open sets are open

A set X endowed with the topology $\mathcal{T}$ is called a **topological space**.

In addition to open sets, sets belonging to our topological space, we can also have closed sets defined as the compliments of open sets. More formally, given open set $U \subset X$ then the compliment of U in X, written $X - U$, is a closed set. It is important to note that open and closed sets do not work like the English interpretation of those words imply: Topological sets are not like doors, they can be open, closed, both, or neither. In addition to the basic definition of a topology we will also need to define a Hausdorff space.

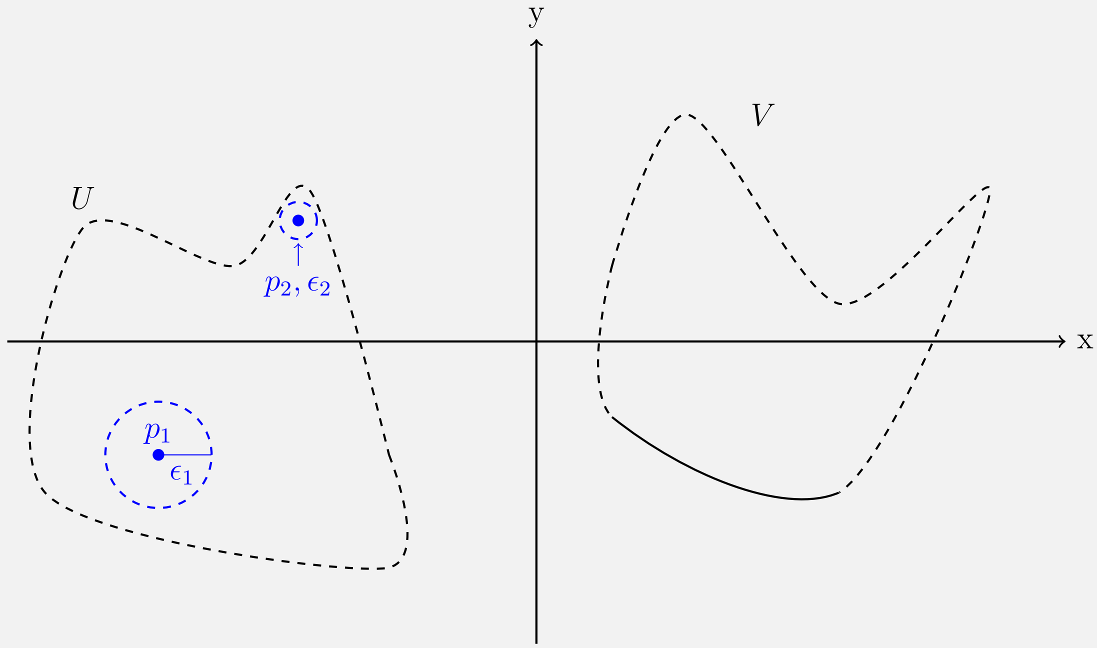
Definition 4.1.2: Hausdorff

A **Hausdorff** space, X, is a topological space in which, for every $x, y \in X$ with $x \neq y$, there exist open sets $U, V \subset X$ such that $x \in U$, $y \in V$, and $U \cap V = \emptyset$. Less formally this means that any two points can be separated by disjoint open neighborhoods. Hausdorff imposes an intrinsic notion of separation between points in a set.

At this point we should probably look at an example:

Example 4.1.1

Consider $\mathbb{R}^n$ with the standard topology that is inherited from its structure as a metric space. A set $U \subseteq \mathbb{R}^n$ is open if and only if for every $p \in U$ there is some $\epsilon > 0$ such that the ball $B_\epsilon(p) = \{x \in \mathbb{R}^n \mid \text{distance}(x, p) < \epsilon\} \subseteq U$. In other words, for every point p in U there is some neighborhood of points around p that are also in U. If we want to understand this more intuitively let us look at $\mathbb{R}^2$:



Note that U is an open set because every point we choose has an open neighborhood (p_1 and p_2 are included to illustrate this) but V is not open because it includes a part of the boundary and none of those points can have an open neighborhood completely in V.

For a much more in depth discussion of topology see [M]. In addition to the extra details and background provided therein, Munkres also includes numerous excellent diagrams to illustrate topological concepts.

4.2 Topological Manifolds

With this basis in topology we can move on to discussing manifolds using our knowledge of topology to guide us. Before defining a topological manifold we need to first define “locally Euclidean”.

Definition 4.2.1: Locally Euclidean

A space X is **locally Euclidean** if every point in the space has an open neighborhood (i.e., an open set containing that point) surrounding it and that neighborhood is homeomorphic to an open subset of $\mathbb{R}^n$ for some fixed n. This is not entirely clear without also knowing the definition of homeomorphic so for clarity: Take two topological spaces A and B with a bijective function $f : A \to B$. If both $f : A \to B$ and $f^{-1} : B \to A$ are continuous then f is a **homeomorphism** between A and B. We could also say that A and B are homeomorphic spaces. It should also be noted, since this will be important later, that homeomorphisms preserve all topological properties.
Note for clarity: A mapping $f : X \to Y$ is continuous if the inverse image of an open set is also open, that is, given an open set $V \subseteq Y$ then $f^{-1}(V) = U \subset X$ is open in X.

Let us take a moment to consider the consequences of a space being locally Euclidean. The definition of locally Euclidean says that a space is locally Euclidean if at any point we choose to look at in the space, when we "zoom in" far enough it will be totally indistinguishable from a point simply sitting in $\mathbb{R}^n$.

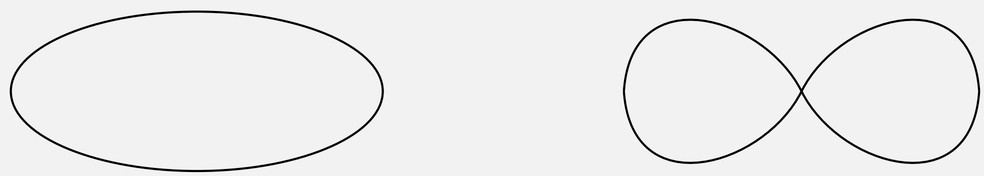
Definition 4.2.2: Topological Manifold

A **topological manifold** is a set X that is a locally Euclidean, Hausdorff space. This means that for every $p, q \in X$ there exist open subsets P and Q of X such that $p \in P$, $q \in Q$, and $P \cap Q = \emptyset$. Additionally, the continuous, invertible mapping f exists such that $f : U \to f(U) \subset \mathbb{R}^n$ for some open subset $U \subseteq X$ and f^{-1} is continuous.

Included below are a couple of examples and/or non-examples of topological manifolds.

Example 4.2.1

Examples (a): Topological manifolds can be simple closed curves like a circle or ellipse since these are locally Euclidean at every point, homeomorphic to $\mathbb{R}$.
Example (b): On the other hand a non-example could be something like a figure-8 or an infinity symbol which each have a point, the crossing, at which it does not behave like a copy of $\mathbb{R}^n$.



✓ Topological Manifold $\qquad$ X Topological Manifold

Example (c): Another example of a topological manifold, this time two dimensional, would be a cone. We can see the homeomorphism between the cone $\mathbb{R}^2$ by simply projecting it directly onto a plane that is normal to the axis of the cone. Note that although this is a topological manifold it is not a smooth manifold since the vertex is not differentiable. This will be discussed in more detail in the next section.

While these are useful definitions and notions of manifolds, we will need something stronger if we want to apply the tools of calculus to manifolds. For a much more detailed discussion of topological manifolds see [LEE].

4.3 Differentiable Manifolds

In this section we will explore manifolds on which we can use calculus and start getting some powerful results. First though we need to define a smooth manifold and in order to do that we will need a couple more definitions.

Definition 4.3.1: Smooth Map

A function $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is **smooth**, also notated "of class $\mathbf{C}^\infty$", if it is infinitely differentiable in each component. This could be rephrased as f is smooth if

$$\frac{\partial^k f}{\partial x^{i_1} \cdots \partial x^{i_k}}$$

exists and is continuous for all k and every combination of $(i_1, \cdots, i_k)$.

Definition 4.3.2: Diffeomorphism

A function f is a **diffeomorphism** if it is a homeomorphism with the added condition that both f and f^{-1} are smooth. Note that this will be necessary when we go to do calculus later.

A good discussion of smooth maps and diffeomorphisms can be found in J. Lee's "Manifolds and Differential Geometry" section 1.4 [LEE]. With these definitions we can now explore what it means to be a smooth manifold.

Definition 4.3.3: Smooth Manifolds

Similarly to the notion of a topological manifold is that of a (**smooth**) **manifold**. Such a manifold, $\mathcal{M}$, is a Hausdorff topological space and contains a countable collection of **patches**, open sets $\{U_i\}$ that cover the manifold. Additionally, there must be a set of coordinate maps $\{\phi_i\}$ satisfying the following two conditions.

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

2. *Compatible on overlaps*: If U_i and U_j are two overlapping coordinate patches, that is $U_i \cap U_j \neq \emptyset$, with associated coordinate maps ϕ_i and ϕ_j then the mapping $\phi_j \circ \phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$ is a diffeomorphism.

Note that if we replaced diffeomorphism with homeomorphism in this definition then we would fall back on a definition for a topological manifold. This illustrates the important step we took in defining a smooth manifold: Just being a manifold is not sufficient, we also need the manifold to be smooth or differentiable so that it works well with calculus. Before proceeding to discuss this definition it could be helpful to have a picture of what is going on (see Figure 1 below).

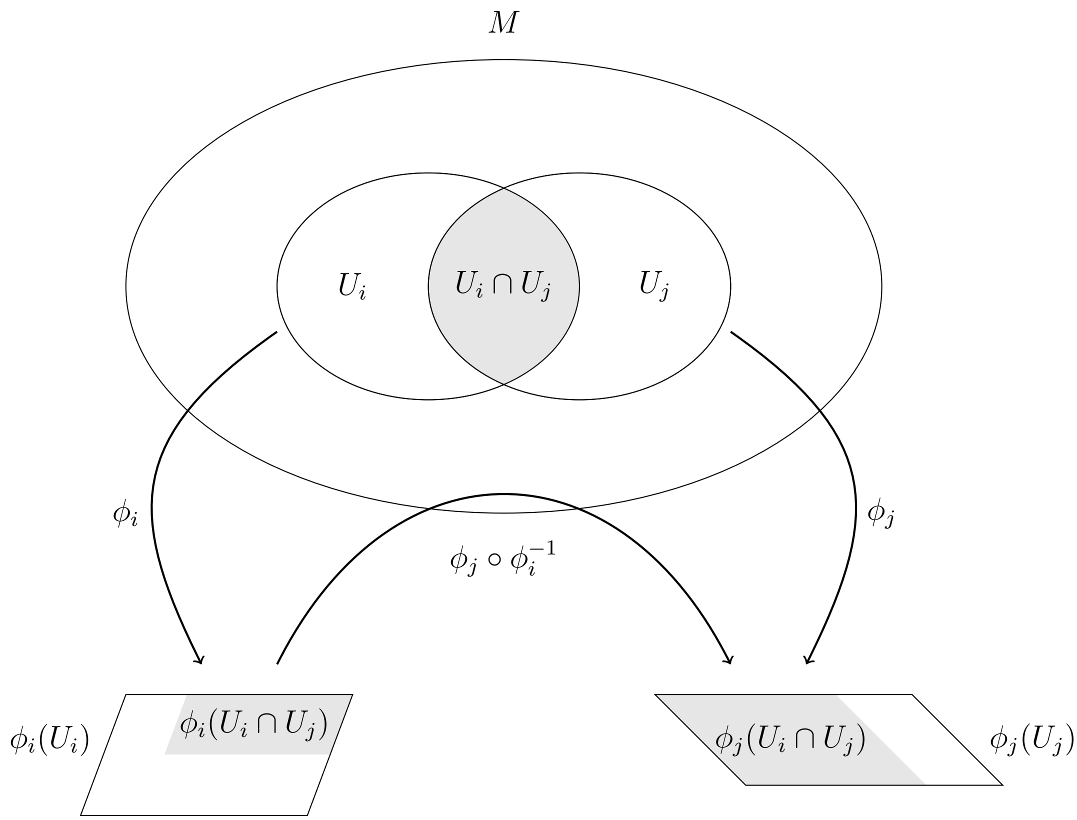


Figure 1: The function that maps between $\phi_i(U_i \cap U_j)$ and $\phi_j(U_i \cap U_j)$ is a diffeomorphism. This shows that regardless of what happens to these sets when they get mapped from M to $\mathbb{R}^n$ their overlap must still act correctly and map easily from one to the other.

We should probably also discuss how we can map between manifolds.

Definition 4.3.4: Smooth Mapping Between Manifolds

Consider manifolds M of dimension m and N of dimension n. The mapping $F : M \to N$ is a smooth mapping between manifolds if F is smooth when represented in coordinates. Specifically, given patches $\phi : U \to \phi(U)$ on M and $\psi : V \to \psi(V)$ on N with $F(U) \cap V \neq \emptyset$ we have that

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

is smooth and maps from a subset of $\mathbb{R}^m$ to a subset of $\mathbb{R}^n$.

As a result of the definition of a smooth manifold we can see that the identity map for such a manifold $\mathrm{id} : M \to M$ must also be smooth. In addition to the definitions given above there are a couple more terms we should define. When we pair a patch with a coordinate map we get a **coordinate chart** such as (U_i, ϕ_i). Also, using an exceptionally appropriate notation we say that the collection of all coordinate charts is called an **atlas**. To better understand these concepts we should go over a few examples.

Example 4.3.1

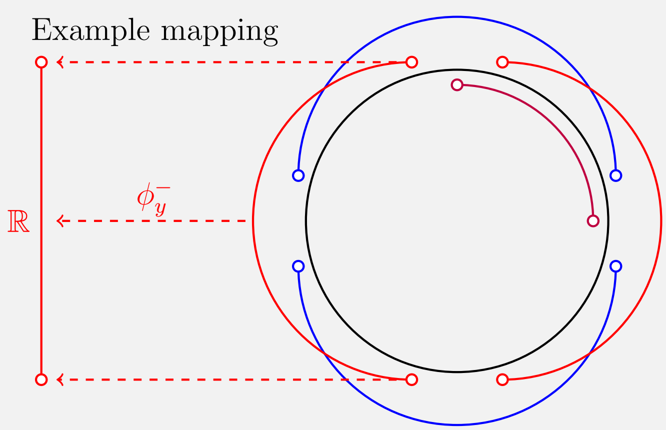
Consider the manifold $\mathbb{R}^2$. Note that this is a manifold because $\mathbb{R}^2$ is (assuming as is common that we are using the standard topology) a Hausdorff topological space, it is clearly locally Euclidean everywhere, and it is compatible on overlaps. We can cover this manifold with only one patch, $\mathbb{R}^2$ itself, and use the identity coordinate map. That means the coordinate chart $(\mathbb{R}, \mathrm{id}_{\mathbb{R}^2})$ is all we need to describe this manifold.

Example 4.3.2

In this example we want to prove that $S^1 = \{(x,y)|x^2+y^2=1\}$, the unit circle, is a smooth manifold. To do this we need to show that it satisfies all of the properties in our definition of a smooth manifold. We can fairly easily see that this is both Hausdorff and locally Euclidean because each open subset of S^1 can be mapped onto an open subset of $\mathbb{R}$. Showing that it is compatible on overlaps is slightly more difficult and we will need to define the coordinate charts first. Although there are more efficient ways to assign coordinate charts[i] in this example we will use the four charts defined below:

$$\phi_x^+ : U_x^+ = \{(x,y)|x^2+y^2=1, y>0\} \to \mathbb{R} \qquad \text{as} \qquad \phi_x^+(x,y) = x$$
$$\phi_x^- : U_x^- = \{(x,y)|x^2+y^2=1, y<0\} \to \mathbb{R} \qquad \text{as} \qquad \phi_x^-(x,y) = x$$
$$\phi_y^+ : U_y^+ = \{(x,y)|x^2+y^2=1, x>0\} \to \mathbb{R} \qquad \text{as} \qquad \phi_y^+(x,y) = y$$
$$\phi_y^- : U_y^- = \{(x,y)|x^2+y^2=1, x<0\} \to \mathbb{R} \qquad \text{as} \qquad \phi_y^-(x,y) = y$$

These will map the top, bottom, left, and right halves (excluding the endpoints of each arc) onto the axes showing a homeomorphism between these chunks and sections of the real line. A diagram of these sections (in a semi-"exploded" view) is provided for reference:



The red curves are the positive and negative halves of the circle that get mapped onto the y-axis while the blue curves are the same for the x-axis. The purple curve shows the overlap of ϕ_x^+ and ϕ_y^+.

In order to check that this is compatible on overlaps we need to check each pair of mappings that overlap although we will only show one of the four since they are very repetitive. For the overlap $U_x^+ \cap U_y^+ = \{(x,y)|x^2+y^2=1, x>0, y>0\}$ we will use $(\phi_x^+) \circ (\phi_y^+)^{-1} : \phi_y^+(U_x^+ \cap U_y^+) \to \phi_x^+(U_x^+ \cap U_y^+)$ to map between these and show that it is a diffeomorphism.

$$(\phi_x^+) \circ (\phi_y^+)^{-1}(y) = \phi_x^+\left(\sqrt{1-x^2}, y\right) = \sqrt{1-x^2}$$

Note that, given the domain restrictions inherent to this overlap, we have found a perfectly valid diffeomorphism meaning this overlap is compatible. We get the same result with varying combinations of negatives for the other three overlaps.

One type of example we have not looked at yet is what happens if we have higher dimensional manifolds. Given the detail in the last example it should be sufficient to merely sketch out the process for a higher dimensional manifold without going into too much detail.

Example 4.3.3

Consider the unit sphere $S^2 = \{(x,y,z)|x^2+y^2+z^2=1\}$. As with the previous example we can see that this set is Hausdorff and locally Euclidean. If we define coordinate charts in a similar way to the previous example we get very similar results although this time we need six coordinate charts instead of four. In this case we are projecting half spheres onto the x, y, and z planes as opposed to projecting half circles onto the x and y axes.

$$\phi^+_{xy} : U^+_{xy} = \{(x,y,z)|x^2+y^2+z^2=1, z>0\} \to \mathbb{R}^2 \qquad \text{as} \qquad \phi^+_{xy}(x,y,z) = (x,y)$$

$$\vdots \qquad\qquad\qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots$$

$$\phi^-_{yz} : U^-_{yz} = \{(x,y,z)|x^2+y^2+z^2=1, x<0\} \to \mathbb{R}^2 \qquad \text{as} \qquad \phi^-_{yz}(x,y,z) = (y,z)$$

We will check one of the 12 overlaps to verify that compatibility holds. Arbitrarily let us choose ϕ^+_{xy} and ϕ^-_{yz} which overlap on the region $U^+_{xy} \cap U^-_{yz} = \{(x,y,z)|x^2 + y^2+z^2=1, z>0, x<0\}$.

$$\left(\phi^+_{xy}\right) \circ \left(\phi^-_{yz}\right)^{-1} : \phi^-_{yz}(U^+_{xy} \cap U^-_{yz}) \to \phi^+_{xy}(U^+_{xy} \cap U^-_{yz})$$

$$\downarrow$$

$$\left(\phi^+_{xy}\right) \circ \left(\phi^-_{yz}\right)^{-1}(y,z) = \phi^+_{xy}\left(-\sqrt{1-y^2-z^2}, y, z\right) = \left(-\sqrt{1-y^2-z^2}, y\right)$$

Note that while this is restricted to the specified domain it is a diffeomorphism so these overlaps can work together although this is not quite all. In a little while we will discuss compatibility of orientation and show that these charts just need one small tweak to be totally correct.

Although this does allow us to check that something is a manifold, as should be apparent from these last example this process can get long and tedious so it would be nice if there was a better way. Fortunately there is a better way to do this, but we will need a handful of extra definitions to do it. First note that when we talk about local coordinates we mean coordinated defined by a diffeomorphism f such that the following holds.

$$f : U \to f(U) \subset \mathbb{R}^m \qquad \text{where} \qquad p \mapsto (x^1, \ldots, x^m)$$

[i] In fact, we could use just two coordinate charts each just leaving out one point. If we are only concerned with integration over our manifold, we could actually use just one coordinate chart. This is because points are boundaries for a 1-dimensional manifold and as will be discussed later the boundary, since it is lower dimensional, does not contribute to integrals and can for those purposes be omitted.

What this basically says is that we are defining coordinates in a locally Euclidean patch near some point.

Definition 4.3.5: Immersions and Embeddings

Consider smooth manifolds M with dimension m and N with dimension n. Let $f : M \to N$ be a smooth map represented in local coordinates as:

$$f(x^1, \cdots, x^m) = (f^1(x^1, \cdots, x^m), \cdots, f^n(x^1, \cdots, x^m))$$

Using bad notation this will often be written as:

$$f(x^1, \cdots, x^m) = (y^1(x^1, \cdots, x^m), \cdots, y^n(x^1, \cdots, x^m)) = (y^1, \cdots, y^n)$$

Now if the Jacobian matrix of this transformation, $Df(x) = (\partial f^i/\partial x^j)$ or just $(\partial y^i/\partial x^j)$ using sloppy notation, has maximal rank at $p \in M$ there are two cases.

1. If $m \leq n$ then the rank of the Jacobian is m and f is called an **immersion at** p
2. If $n \leq m$ then the rank of the Jacobian is n and f is called a **submersion at** p

Additionally, we have a couple other definitions that follow this including: If f is an immersion for all $p \in M$ then M is an immersed **submanifold** of N. If f is an immersion and it is injective then it is called an **embedding**.

These definitions lead straight into two very important theorems, the first of which is the Whitney embedding theorem, see [R] or for a much more nuanced and deep discussion see [Sk] section 2.

The Whitney Embedding Theorem

The Whitney embedding theorem states that any n-dimensional topological manifold can be embedded in $\mathbb{R}^{2n+1}$ and any n-dimensional smooth manifold can be embedded in $\mathbb{R}^{2n}$.

One of the main positives to this theorem is that it allows us to embed a complicated manifold in some $\mathbb{R}^k$ space that is more familiar and easier to work in. Before moving on to the next big theorem we should define a couple more terms.

Definition 4.3.6: Regular and Critical Points

Suppose we have a mapping $f : M \to N$ where $m \geq n$. If f is a submersion at $p \in M$ then p is a **regular point** of f. If p is not a regular point of f then we call it a **critical point**. Given some point $q \in N$, if every point in $f^{-1}(q)$, the **fiber** over q, is a regular point (meaning any point that maps to q must be regular) then we call q a **regular value**.

The Regular Value Theorem

Again consider the map $f : M \to N$ between manifolds of dimension m and n respectively. Let $q \in N$ be a regular value of f. Then $f^{-1}(q)$ is a smooth embedded submanifold of M with dimension $m - n$.
(See [R] for a proof)

With this theorem we can now easily determine whether something is a manifold or not. To show this let us revisit the example where we attempted to show that a circle is a manifold.

Example 4.3.4

In order to apply this theorem we will need a mapping between manifolds. Since we know that all $\mathbb{R}^n$ are manifolds let us define our mapping as $f : \mathbb{R}^2 \to \mathbb{R}$ where $f(x, y) = x^2 + y^2$. Note that the Jacobian matrix, $Df = \langle 2x, 2y\rangle$, has the maximal possible rank of 1 and can not be zero unless $(x, y) = (0, 0)$. This implies that the mapping is a submersion regardless of the point we choose. Now let us look at the fiber over 1: $f^{-1}(1) = \{(x, y)|x^2 + y^2 = 1\}$. As stated previously each $p \in \mathbb{R}^2$ must be a regular point therefore every point in $f^{-1}(1)$ must be regular so $1 \in N$ is a regular value. Finally, by the regular value theorem it must then be the case that $f^{-1}(1) = \{(x, y)|x^2 + y^2 = 1\} = S^1$ must be a smooth embedded submanifold of $\mathbb{R}^2$ with dimension 1.

And that was clearly much easier than previous methods of establishing if something is a manifold. This method can be used on many of the objects we want to study and its simplicity should be admired since at the heart of the matter all we really did was check that the Jacobian matrix behaved correctly then we got, almost for free, the result that the level surface we wanted to study was a manifold.

5 Tangents

Now that we have some basic understanding of manifolds one of the next logical directions to proceed if we want to end up doing calculus is to study tangents to manifolds. Not just tangents though, we also want to study the spaces they make up, the bundles formed by all of the tangent spaces, dual spaces, and how these tangent ideas build towards our next topics.

5.1 Tangent Spaces

In order to study tangents we will need to understand tangent spaces. The first thing to recognize about tangent spaces is that we can think about them from multiple useful perspectives. The first, and less formal, perspective we will consider is heavily based on coordinates to define a basis for the tangent space. The main definition will be given in part 2 when we talk about the more "official" perspective on tangent spaces.

1. Let (ϕ, U) be a patch on manifold M with point $p \in U$.

   $$\text{Assume } \phi : U \to \mathbb{R}^n \qquad \text{where} \qquad \phi(p) = (x^1(p), \cdots, x^n(p))$$

   Then we will define a basis set for the tangent space at p as the set:

   $$\left\{ \left.\frac{\partial}{\partial x^1}\right|_p, \cdots, \left.\frac{\partial}{\partial x^m}\right|_p \right\} \qquad \text{where} \qquad \text{span}\left\{ \left.\frac{\partial}{\partial x^1}\right|_p, \cdots, \left.\frac{\partial}{\partial x^m}\right|_p \right\} = T_pM$$

   And any $X_p \in T_pM$ can be written as:

   $$X_p = \sum_i X^i \left.\frac{\partial}{\partial x^i}\right|_p$$

   Where each X^i is a scalar and the $\partial/\partial x^i$ component of X_p. Further, if we have another valid basis for T_pM using some coordinates $\partial/\partial y^j$ then we can use the chain rule to give us:

   $$\frac{\partial}{\partial y^j} = \sum_i \frac{\partial x^i}{\partial y^j}\frac{\partial}{\partial x^i} \qquad \to \qquad Y^j = \sum_i X^i \left.\frac{\partial y^j}{\partial x^i}\right|_p$$

   Each $\partial y^j/\partial x^i$ is given by the change of coordinates matrix. These more coordinate heavy notations will be more useful later when we discuss the dual space of a tangent space.

2. Before we get to dual spaces let us switch to talking about the second perspective and more formal definition of tangent spaces. This perspective is based on algebraic properties and it requires the introduce of linear derivations. Note that $\Omega^0(M) = \{f : M \to \mathbb{R} \,|\, f \text{ is smooth}\}$ is the space of smooth scalar valued functions, these are sometimes called test functions.

Definition 5.1.1: Tangent Spaces

Let M be a manifold with dimension m. The **tangent space to M at p** is a vector space written as T_pM. The elements of this space, X_p, are called **tangent vectors** and $X_p \in T_pM$ if and only if X_p is a **linear derivation** at p. Linear derivations essentially capture our basic intuitions about derivatives but in a more general way. To elaborate on this take $a, b \in \mathbb{F}$ and $f, g \in \Omega^0(M)$ then $X_p : \Omega^0(M) \to \mathbb{R}$ satisfies the following conditions.

1. Linearity: $X_p(af + bg) = aX_p(f) + bX_p(g)$
2. The Leibniz property: $X_p(fg) = g(p)X_p(f) + f(p)X_p(g)$

Notice that, given a patch (U, ϕ), if $f, g \in \Omega^0(M)$, then

$$\frac{\partial}{\partial x^i}\bigg|_p \Big[f\Big] = \frac{\partial}{\partial x^i}\Big[f \circ \phi^{-1}(x^1, \ldots, x^n)\Big]\bigg|_{\phi(p)}$$

In other words, $\frac{\partial}{\partial x^i}\big|_p (f)$ is the partial derivative of the coordinate version of f with respect to the i-th coordinate and evaluated at the coordinates of the point p. In particular, this is linear and

$$\frac{\partial}{\partial x^i}\bigg|_p \Big[fg\Big] = \frac{\partial f}{\partial x^i}\bigg|_p \cdot g(p) + f(p) \cdot \frac{\partial g}{\partial x^i}\bigg|_p$$

the Leibniz property holds by the standard product rule. Thus $\frac{\partial}{\partial x^i}\big|_p$ is in fact a linear derivation at p. In other words, our tentative tangent vector at p is an official tangent vector at p.

3. For now let us again switch directions and consider a third approach to viewing tangent spaces. This last view of tangent spaces in more geometric in nature and revolves around equivalence classes of curves. Let α and β be two smooth mappings from $(-1, 1)$ to M. Also, let $U, V \subset M$ with charts ϕ mapping to x^i coordinates and χ mapping to y^i coordinates such that $\alpha(0) = \beta(0) = p \in U \cap V$. An interesting definition of the type of equivalence class we are attempting to understand is given in [LEE]. He states that an equivalence relation on the set Γ_p of all triples[ii] $(p, v, (U, x))$ can be defined by requiring that $(p, v, (U, x)) \sim (p, w, (V, y))$ if and only if

$$w = D\left(y \circ x^{-1}\right)\big|_{x(p)} \cdot v$$

[ii] This notation is a bit different from what has been used previously in this paper so I will break it apart a little: This triple tells us that we are dealing with a tangent v at point p represented in the coordinates from the chart (U, x) which takes the set $U \subset M$ and maps it as $x : M \to \mathbb{R}^n$

Which means that two tangent vectors, v and w, are equivalent if they can be related by the derivative at $x(p)$ of the coordinate change $y \circ x^{-1}$ between the coordinates of the two charts. More discussion on tangent spaces can be found in [R] chapter 3. Basically, in this third viewpoint, tangents at p are seen as equivalence classes of curves on our manifold that pass through p. This equivalence codifies the idea of curves being tangent to each other.

5.2 The Tangent Bundle

At this point one of the next logical questions is, since we have a tangent space at every point in the manifold, can we put these all together in some meaningful way? The answer is yes, the collection of all tangent vectors at every point in M gives us the **tangent bundle**. If the dimension of M is m then objects in the tangent bundle can be described by $2m$-tuples. Pick a chart (U, ϕ) where $\phi(p) = (x^1(p), \ldots, x^m(p))$ and the x^i are coordinate functions for ϕ. Then if $X_p \in T_pM$, we have

$$X_p = X^i \left. \frac{\partial}{\partial x^i} \right|_p$$

Thus we can encode X_p in coordinates via $(x^1(p), \ldots, x^m(p), X^1, \ldots, X^m)$. Conversely, any $2m$-tuple whose first m coordinates are given by $\phi(p)$ and last m coordinates come from a random vector $\mathbf{v} = \langle v^1, \ldots, v^m \rangle$ *i.e.* $(\phi(p), \mathbf{v}) = (x^1(p), \ldots, x^m(p), v^1, \ldots, v^m)$, defines an element of T_pM namely

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

More formally we can define tangent bundles as follows:

Definition 5.2.1: Tangent Bundle

If we have a smooth manifold M and a tangent space T_pM at every point $p \in M$ then we can build the **tangent bundle**, denoted TM, by taking the disjoint union of all possible tangent spaces or by unioning the tangent spaces at each distinct point.

$$TM = \dot{\bigcup_{p \in M}} T_pM = \bigcup_{p \in M} \{p\} \times T_pM$$

Now consider the nature of this space. The first m coordinates are from $p \in U \subset M$ but recognize that we have a smooth manifold so this must be locally Euclidean hence $U \subset \mathbb{R}^m$. This means that $p \in \mathbb{R}^m$. We also have $\mathbf{v} \in \mathbb{R}^m$ so a coordinate patch of this tangent bundle looks like $(U \subset \mathbb{R}^m) \times \mathbb{R}^m \subset \mathbb{R}^{2m}$ and it can be shown (see [F]) that this is a $2m-$dimensional differentiable manifold. Since the tangent bundle is a manifold there are a few interesting things about it that we should investigate.

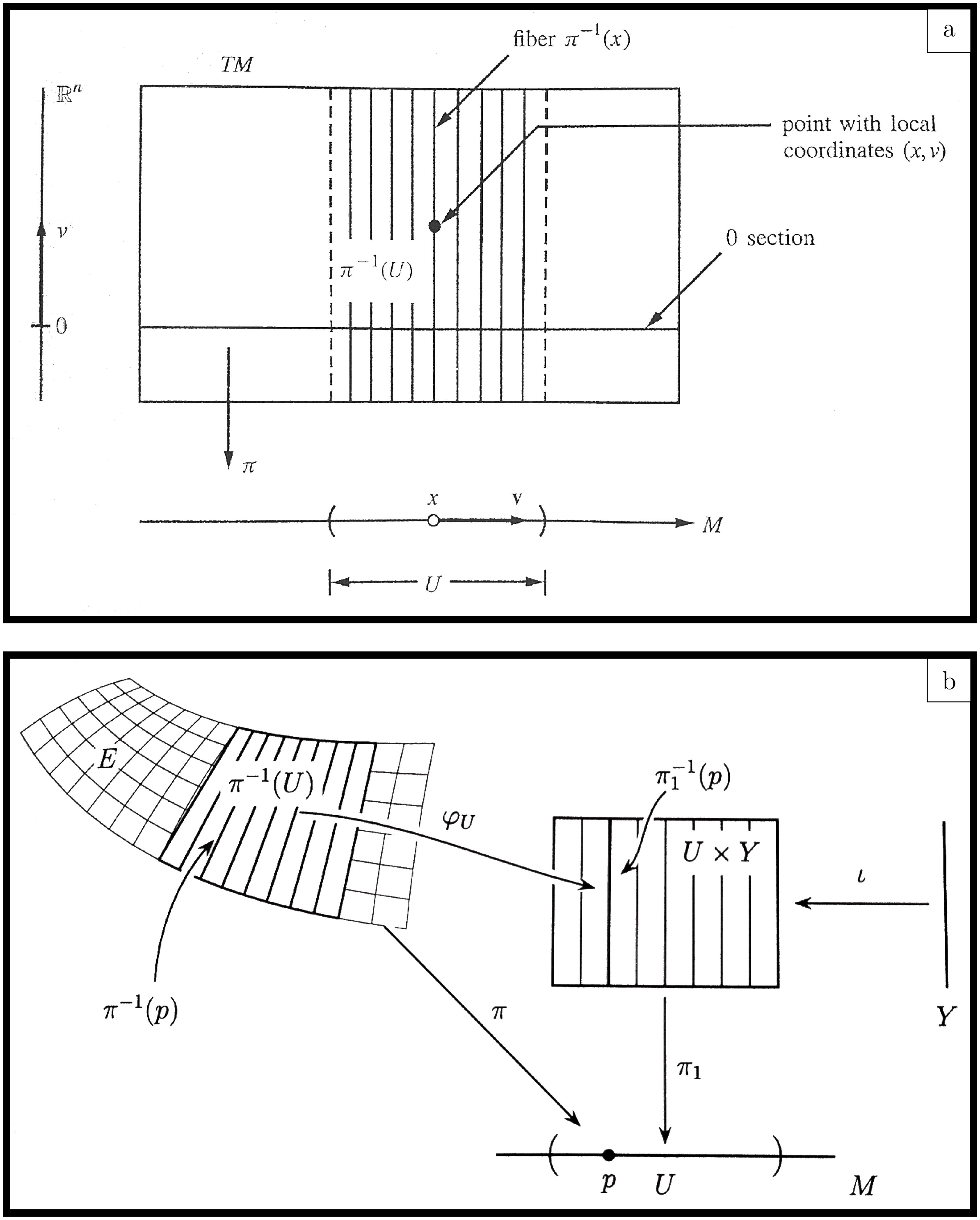


Figure 2: (a) is a diagram from [F] showing the tangent bundle and some of the ways in which we can interpret its facets. (b) is a similar diagram from [R] using slightly different notation (although this is actually a diagram from the vector bundle section).

First, it makes sense to define a **projection** back onto our original manifold.

$$\pi : TM \to M \qquad \text{defined as} \qquad \pi(p, \mathbf{v}) = p$$

Note: One can show that π is a smooth map from TM to M. For convenience of notation such a bundle is often written as a 4-tuple: $(TM, M, \mathbb{R}^m, \pi)$. This just summarizes its important characteristics in a concise way. If we instead go backwards from this projection mapping we get $\pi^{-1}(p) = \{(p, X_p) \in TM \mid \pi(p, X_p) = p\} = T_pM$. This is a copy of $\mathbb{R}^m$, which we call the **fiber** over p. Suppose one selects a vector X_p for each $p \in M$ and does so in a smooth way. In other words, $X : M \to TM$ where $p \mapsto (p, X_p)$ is a smooth map between M and TM (although we will identify $\{p\} \times T_pM$ with T_pM hereafter for ease of notation). We call such maps **vector fields**. Notice that $\pi \circ X = \mathrm{id}_M$ since X assigns a point a vector and then π immediately removes the vector taking us back to the original point in M. This is called a **section** of the tangent bundle TM. Figure 2 shows graphical/geometric interpretations of some of these concepts from both [F] and [R].

5.3 Cotangent Spaces and More Bundles

Just as we discussed the dual to a vector space earlier, we can also have the dual to a tangent space. The dual space to T_pM is denoted T_p^*M and is usually called the cotangent space at p. The elements of the cotangent space are called cotangent vectors and, similarly to what dual vectors did, these cotangent vectors provide a linear mapping between T_p^*M and $\mathbb{R}$. Also, the bases behave in a way comparable to the vector/dual vector space bases. That is, given the basis

$$\left\{ \left.\frac{\partial}{\partial x^i}\right|_p \;\middle|\; i = 1, \ldots, m \right\}$$

for T_pM we can define the (dual) basis for T_p^*M as $\{d_px^i\}$ where

$$\left\langle d_px^j, \left.\frac{\partial}{\partial x^i}\right|_p \right\rangle = \left\langle \left.\frac{\partial}{\partial x^i}\right|_p, d_px^j \right\rangle = d_px^j\left(\left.\frac{\partial}{\partial x^i}\right|_p \right) = \delta_i^j$$

Recall that given $f \in V^*$ and $v \in V$, $\langle f, v\rangle = \langle v, f\rangle = f(v)$ is the dual pairing between V and V^* and just means that we plug v into the dual vector, a scalar valued linear map, $f : V \to \mathbb{R}$. This is exactly what we used when discussing dual vector spaces, the elements of the dual basis pick out the corresponding element of the regular basis and leave everything else as zero. Note that these basis elements are differential operators and so we will call a general element of this space a **differential form**. In coordinates, $\alpha_p \in T_p^*M$ is $\alpha_p = a_i(p)\, d_px^i$. Suppose we smoothly assign a dual vector to each point in our manifold. Say, $\alpha : M \to T^*M$ is smooth where $\alpha(p) = \alpha_p \in T^*M$ and T^*M is the cotangent bundle defined much like the tangent bundle. Then α is a section of the cotangent bundle and we call it a differential 1-form. Although it won't be discussed in any detail yet we define a **differential k-form** as a smooth assignment of an element of

$\bigwedge^k(T_p^*M)$ to each $p \in M$.
To finish off this section I want to talk about bundles just a little more. First, since we just discussed cotangent spaces we should mention cotangent bundles. Due to being very similar to tangent bundles we won't say that much about cotangent bundles, but here is a basic sketch. We can essentially change the definitions for a tangent bundle to use cotangent vectors and cotangent spaces to arrive at a reasonable definition of cotangent bundles. Something like:

Definition 5.3.1: Cotangent Bundle

If we have a smooth manifold M and a cotangent space T_p^*M at every point $p \in M$ then we can build the **cotangent bundle**, denoted T^*M, by taking the disjoint union of all possible cotangent spaces or by unioning the cotangent spaces at each distinct point.

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

All of the other definitions translate accordingly including projections, fibers, sections, etc. Now let us move on for now and discuss general vector bundles.
Similarly to tangent bundles and cotangent bundles, we define a vector bundle as follows using the notations from [R].

Definition 5.3.2: Vector Bundle

A **vector bundle**, E, over a manifold M has the projection map $\pi : E \to M$ and must satisfy three axioms:

1. *Fiber isomorphism*: Each fiber $\pi^{-1}(p)$ with $p \in M$ is isomorphic to the vector space Y of dimension m.

2. *E is a product (locally)*: Consider the manifold E. For all $p \in M$ there is some neighborhood U containing p and a diffeomorphism

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

   Where Y is the fixed vector space from the first axiom. This mapping is called a **local trivialization** of E over U.

3. *ϕ_U carries fibers to fibers linearly*: Noting that $\phi_U : \pi^{-1}(p) \to \pi_1^{-1}(p)$ is linear for all p, we can also say that the map $\pi_1 : U \times Y \to U$ is a projection onto the first element such that $(p, f) \mapsto p$.

As with previous mentions of bundles we will denote a vector bundle with the 4-tuple (E, M, Y, π), notice also that E is a manifold as it (the bundle space) was for tangent bundles. See Figure 2b for the corresponding diagram from [R].

Vector bundles smoothly assign a vector space to every point in M. We can now see that the tangent bundle and cotangent bundle were special cases of vector bundles where these vector spaces were the tangent or cotangent space at every point. In the interest of time we won't spend any longer on vector bundles, although [R] has a good discussion of them in chapter 7. For now let us continue by talking about a generalization of both vector and covector fields, namely tensor fields. These will yield a new family of examples of vector bundles.

5.4 Tensor Fields

Recall from our first discussion of tensors that a tensor can be expressed as a multilinear map or as an element of a tensor product space. In this section we will revisit both of these ideas and connect them to some of the concepts we discussed concerning tangent spaces and cotangent spaces. At its heart a **tensor field** is simply a smooth assignment of a tensor to every point of M. The complexity comes when we try to express the tensors since the choice of how to express the tensors, along with actually expressing them, has implications on the tensor field.

Tensor product space
The tensor product space interpretation, as discussed in the tensor section, is much more coordinate heavy so we will start by giving the standard setup: Let M be a manifold with patch U using local coordinates $x^1, x^2, \cdots, x^m$. For any point $p \in U$ we can consider the vector fields $\partial/\partial x^i$ and 1-form fields dx^i as being the bases of the tangent space and cotangent space at that point. Putting everything together in a similar way to the general tensors discussed earlier we can create a basis (valid when working with points $p \in U$) for the tensor field from basis elements of the form:

$$\frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}$$

These form a basis for the tensor product space $(T_pM)^{\otimes r} \otimes (T_p^*M)^{\otimes s}$ where the elements look (locally in coordinates) like:

$$\Psi = \Psi^{i_1 \cdots i_r}_{j_1 \cdots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}$$

Just as we had "tensors of type (r, s)" in this case we have a **tensor field Ψ of type (r, s)**.

As multilinear maps
Following along with the discussion in [R] chapter 3 let $\tilde{T}^r_s(p)$ be the space of every multilinear map on:

$$\underbrace{T_p^*M \times \cdots \times T_p^*M}_{r \text{ times}} \times \underbrace{T_pM \times \cdots \times T_pM}_{s \text{ times}}$$

Recall that mapping multilinearly takes objects with some number of "slots" and maps to an object with a different number of "slots". Last time we saw this we mapped n vectors to a single tensor. In this case we are creating a tensor field of type (r, s) by smoothly assigning an element $\Psi_p \in \tilde{T}^r_s(p)$ for each $p \in M$. Note that smooth means for any smooth covector fields $\alpha_1, \cdots, \alpha_r$ and smooth vector fields $X_1, \cdots, X_s$ near p then the map

$$p \mapsto \Psi_p(\alpha_1(p), \cdots, \alpha_r(p), X_1(p), \cdots, X_s(p))$$

is also smooth. Alternatively, smooth means that all component functions $\Psi^{i_1 \cdots i_r}_{j_1 \dots j_s}$ are smooth functions (for an arbitrary choice of coordinates). Upon considering these definitions more carefully we find that not only do we have that Ψ is multilinear, we actually have a stronger case that it is **function linear**. Function linear involves modifying the definition of multilinear, the heart of which given here,

$$\Psi(v_1, \cdots, au + bw, \cdots, v_{r+s}) = a\Psi(v_1, \cdots, u, \cdots, v_{r+s}) + b\Psi(v_1, \cdots, w, \cdots, v_{r+s})$$

to replace a and b, which had previously been scalars, with smooth functions.
Summarizing this:

Definition 5.4.1: Tensor Fields

First, let $\Gamma(TM)$ denote the space of all vector fields on M and likewise $\Gamma(T^*M)$ denotes the space of all covector fields on M. Then a **tensor field Ψ of type (r, s)** is a function multilinear map given by:

$$\Psi : \underbrace{\Gamma(T^*M) \times \cdots \times \Gamma(T^*M)}_{r \text{ times}} \times \underbrace{\Gamma(TM) \times \cdots \times \Gamma(TM)}_{s \text{ times}} \to \mathbb{R}$$

There is a proof of this for the specific case of only vector fields (so $r = 0$) in [R]

Before leaving this section, we can define a differential k-form ω to be an alternating tensor field of type $(0, k)$. In particular,

$$\omega : \underbrace{\Gamma(TM) \times \cdots \times \Gamma(TM)}_{k \text{ times}} \to \mathbb{R}$$

is k-multilinear and alternating meaning $\omega(X_1, \ldots, X_k) = 0$ anytime that $X_i = X_j$ for some $i \neq j$ (i.e., we have a repeated input). This is equivalent to being skew-symmetric (i.e., $\omega(X_1, \ldots, X_k) = (-1)^\sigma \omega(X_{\sigma(1)}, \ldots, X_{\sigma(k)})$ for any permutation σ on $\{1, 2, \ldots, k\}$) and $(-1)^\sigma$ is the sign of σ: $+1$ for even and -1 for odd. Now that we have discussed tensor fields I would like to shift our focus back to manifolds.

6 Manifolds Round Two

In this section we will return to manifolds and explore ideas relating to derivation on manifolds, orientation, manifolds with boundary, and integration on a patch.

6.1 Differentiation

Differential k-forms
Recall from earlier that a differential 1-form is a covector looking like $\alpha = a_i dx^i$, but this is really just a specific k-form. Further, k-forms are just a specific type of tensor: alternating tensors. These tensors are called alternating tensors because of their anti-symmetric property $\alpha_{ij} = -\alpha_{ji}$.

Example 6.1.1

An example is alternating 2-tensors in $\mathbb{R}^3$. The basis of these if given by:

$$\beta = \{dy \otimes dz - dz \otimes dy, dz \otimes dx - dx \otimes dz, dx \otimes dy - dy \otimes dx\}$$

Since these representation are rather cumbersome we can denote them, up to a constant, as:

$$\{dy \wedge dz, dz \wedge dx, dx \wedge dy\}$$

We can use this wedge notation to define our k-forms (on a coordinate patch U at point p) as:

$$\omega = \frac{1}{k!} \sum a_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k} = a_I dx^I$$

Note that $\omega_p \in \bigwedge^k (T_p^* M)$ which are all of the anti-symmetric tensor of type $(0, k)$ Being careful here: $\omega_p \in \bigwedge^k (T_p^* M)$ is an anti-symmetric = alternating tensor of type $(0, k)$. If we collect these for all points and let p vary, then we have a tensor *field*. We could also say that $\omega \in \Omega^k(M)$ which is the vector space of all k-forms.

Example 6.1.2

In the previous example we looked at a basis for alternating 2-tensors on $\mathbb{R}^3$, but here we want to look at general k-forms on $\mathbb{R}^3$.

$\Omega^0(\mathbb{R}^3) = \{f : \mathbb{R}^3 \to \mathbb{R} \mid f \text{ is smooth}\}$
$\Omega^1(\mathbb{R}^3) = \{f\ dx + g\ dy + h\ dz \mid f, g, h \in \Omega^0(\mathbb{R}^3)\}$
$\Omega^2(\mathbb{R}^3) = \{f\ dx \wedge dz + g\ dz \wedge dx + h\ dx \wedge dy \mid f, g, h \in \Omega^0(\mathbb{R}^3)\}$
$\Omega^3(\mathbb{R}^3) = \{f\ dx \wedge dy \wedge dz \mid f \in \Omega^0(\mathbb{R}^3)\}$

Also, $\Omega^4(\mathbb{R}^3) = \{0\}$. Notice how $\Omega^4(\mathbb{R}^3)$ makes sense looking at the way we define the wedge product. Elements of $\Omega^4(\mathbb{R}^3)$ must contain a duplicate term, without loss of generality let it be dx here, and when we go to evaluate this part we get $dx \wedge dx = dx \otimes dx - dx \otimes dx = 0$ hence any k-form where k is greater than the dimension of the space must be 0.

One last thing to note, if we have λ as some m-form then $\deg(\lambda) = m$. With a basic grasp of k-forms now we can ask what does it mean to take the derivative of these?

The exterior derivative

The exterior derivative is a powerful extension of normal derivatives which applies to k-forms. First, the definition:

Definition 6.1.1: The Exterior Derivative

The **exterior derivative** is a linear operator which uniquely maps k-forms to $(k+1)$-forms, $d : \Omega^k(M) \to \Omega^{k+1}(M)$. For any forms λ and μ with $\deg(\lambda) = m$, and any function f the operator d satisfies the following four properties:

1)	Linear:	$d(\lambda + \mu) = d\lambda + d\mu$
2)	Graded derivation:	$d(\lambda \wedge \mu) = d\lambda \wedge \mu + (-1)^{\deg\lambda} \lambda \wedge d\mu$
3)	Nilpotent:	$d^2\lambda = 0$
4)	Natural:	In any local coordinates $\{x^i\}$ about a point p $df = \sum \frac{\partial f}{\partial x^i} dx^i$

From this we can see that it acts similarly to normal derivatives in many ways although it is slightly more complicated. To help clarify the use of the external derivative here is an example on $\mathbb{R}^2$:

Example 6.1.3

Taking a manifold $M = \mathbb{R}^2$ we want to calculate the exterior derivative on a general 1-form $\vec{f} = P\ dx + Q\ dy$:

$d(P\ dx + Q\ dy) = d(P\ dx) + d(Q\ dy)$	Linearity
$= dP \wedge dx + P \wedge d^2x + dQ \wedge dy + Q \wedge d^2y$	Graded derivation
$= dP \wedge dx + dQ \wedge dy$	Nilpotent
$= (P_x\ dx + P_y\ dy) \wedge dx + (Q_x\ dx + Q_y\ dy) \wedge dy$	Evaluate
$= P_x\ dx \wedge dx + P_y\ dy \wedge dx + Q_x\ dx \wedge dy + Q_y\ dy \wedge dy$	Distribute
$= P_y\ dy \wedge dx + Q_x\ dx \wedge dy$	Wedge properties
$= (Q_x - P_y) dx \wedge dy$	Wedge properties
$= \nabla \times \vec{f} = \mathrm{curl}(\vec{f}\ )$	Definitions

Using a similar calculation while working in $\mathbb{R}^3$ and identifying
1-forms: $\omega = P\,dx + Q\,dy + R\,dz$
2-forms: $\eta = P\,dy \wedge dz + Q\,dz \wedge dx + R\,dx \wedge dy$
3-forms: $f\,dx \wedge dy \wedge dz$
with vector fields $\mathbf{F} = \langle P, Q, R \rangle$ and with a scalar valued function f, we get for a scalar valued function df is ∇f (the gradient), $d\omega$ is $\nabla \times \mathbf{F}$ (the curl), and $d\eta$ is $\nabla \cdot \mathbf{F}$ (the divergence).

For more discussion of forms and exterior derivatives see [C]. As interesting as this is though, we still need to discuss some other things about manifolds so we will move on to manifolds with boundaries for now.

6.2 Orientations and Manifolds with Boundary

In this section we will talk about orienting a manifold as will become relevant when we want to integrate over manifolds. We will also tweak the definition of manifolds to allow for them to include the boundary. Previously we had manifolds that did not include "edge parts" that would make it impossible to use open sets to fully cover the manifold, but in this section we must consider what happens if boundary regions are included in the manifold. First though, let us discuss orientation.

Definition 6.2.1: Orientation

Let M be a smooth manifold of $\dim(M) = m$. Two patches are said to be **compatible** if, over the region of their overlap, the Jacobian change of coordinates is positive. A manifold is **orientable** if it can be covered by a set of patches that are all compatible. Alternatively, we define **an orientation** as a non-vanishing ($\omega_p \neq 0$ for all $p \in M$) top form $\omega \in \Omega^m(M)$. Then we say a coordinate chart is compatible with ω if for each point p in a coordinate chart U_i we have

$$\omega_p\left(\left.\frac{\partial}{\partial x^1}\right|_p \cdots \left.\frac{\partial}{\partial x^m}\right|_p\right) > 0$$

One can show that if one has an orientation form ω, then one can find an atlas of compatible charts and vice-versa [LEE].

And now we can define a manifold with boundary after which we will put these ideas together to understand the induced orientation on the boundary. What we will see in the definition of a manifold with boundary is that it is identical to the smooth manifold defined earlier except that it has one small tweak to account for the boundary.

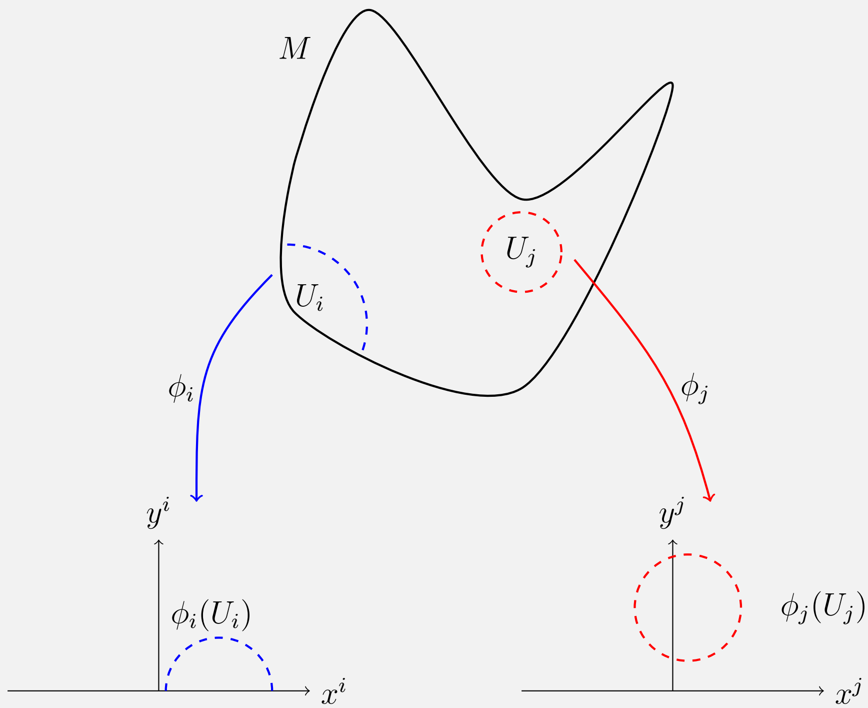
Definition 6.2.2: Manifold with Boundary

A **smooth m-dimensional manifold M with boundary** is a Hausdorff topological space containing a countable collection of coordinate patches $\{U_i\}$ that cover the manifold. Additionally, as we had last time we defined manifolds, when translated to local coordinates using its associated mapping ϕ_i every U_i must be locally Euclidean and every pair of coordinate charts must be compatible on their overlap. Here is the difference for manifolds with boundary: The domains of our charts are allowed to be homeomorphic to open subsets of the upper half space $\mathbb{H}^m = \{(x^1, \cdots, x^m) | x^m \geq 0\}$ instead of $\mathbb{R}^m$.

To illustrate this difference let us consider an example:

Example 6.2.1

Consider a region in $\mathbb{R}^2$ that contains its boundary. If we consider some potential coordinate charts we can start to get a feel for what it means to be homeomorphic to $\mathbb{H}^2$.



Some charts like U_j will remain unchanged, but others like U_i that include the boundary instead get mapped to part of $\mathbb{H}^2$ containing the x-axis boundary.

Ok, but we keep saying "boundary" without really defining it so let us nail down that term before proceeding.

Definition 6.2.3: Boundary

The **boundary** of M, written ∂M, is the set of all points that get mapped to the boundary $\partial\mathbb{H}^m$ of $\mathbb{H}^m$: $\{(x^1, \cdots, x^m)|x^m = 0\}$ in some (and thus every) chart [R]. Using slightly different wording, the boundary of M is the set of all points $p \in M$ such that $\phi_i(p) \in \partial\mathbb{H}^m$. Notice that the boundary of an $m-$dimensional manifold with boundary is itself an $(m-1)$-dimensional manifold *without* boundary.

From this definition we can see that the boundary of each chart in Example 6.2.1 is just a piece of $\mathbb{R}$, this shouldn't be too surprising. Now, as was promised, we will discuss the induced orientation.

Definition 6.2.4: Induced Orientation

To construct the **induced orientation** on the boundary of a manifold M begin with an orientation of the boundary manifold: $\omega \in \Omega^{m-1}(\partial M)$. Now smoothly choose an "outward pointing" normal vector at each point $p \in \partial M$ as follows: $N_p = -\left.\frac{\partial}{\partial x^m}\right|_p \in T_pM$. Finally, construct the induced orientation by using these vectors in the "boundary direction":

$$(v_1, \ldots, v_m) \mapsto \omega_p(N_p, v_1, \ldots, v_{m-1}) = \omega_p\left(-\left.\frac{\partial}{\partial x^m}\right|_p, v_1, \ldots, v_{m-1}\right)$$

for all $p \in \partial M$ and $v_1, \ldots, v_{m-1} \in T_pM$. This gives us an $(m-1)$-form to orient ∂M.

See Figure 3 for a diagram of this as shown in [F].

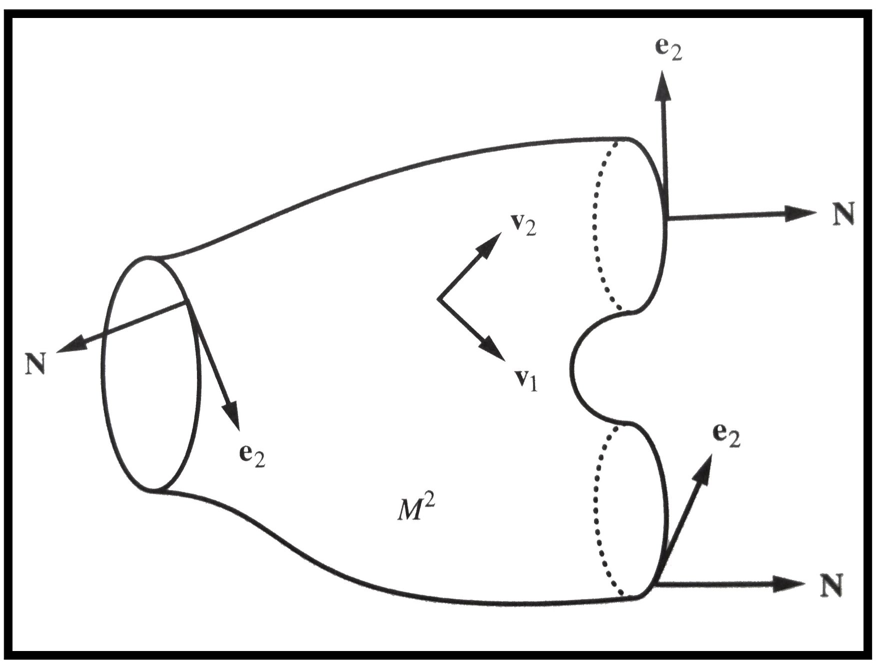


Figure 3: A diagram of a 2-dimensional manifold with boundaries where the outward pointing normal vectors are indicated at various points. These show how we can extend our notion of orientation out to include the boundary.

Now with some understanding of manifolds with boundary let us briefly discuss integration on a patch so that we can move on to discussing the generalized Stokes theorem.

6.3 Calculus on Manifolds

To start with we need to define what it means to integrate over a patch. And just to note this ahead of time, the definition I will give is not quite all there is to it, but it is sufficient for our purposes and the details are readily available in any text that describes integration over a patch in some form [LEE, R, F].

Definition 6.3.1: Integration on a Patch

et M be an m-dimensional manifold and (U, ϕ) be a coordinate chart with local coordinates $x^1, \cdots, x^m$. Further, let $\eta = h\ dx^1 \wedge \cdots \wedge dx^m \in \Omega^m(U)$ be an m-form defined over U (for all $p \in U$). Then the integral of η over U is defined to be:

$$\int_U \eta = \overbrace{\underset{\phi(U)}{\int \dots \int}}^{m \text{ times}} h\ dx^1 \cdots dx^m$$

In other words it works pretty similarly to how we would expect. Changing to another set of coordinates also works similar to our notions of coordinate change from multi variable calculus. Given the notation from the definition above along with another mapping ψ of U to some $\{y^i\}$ coordinates we have:

$$\int_U \eta = \overbrace{\underset{\phi(U)}{\int \dots \int}}^{m \text{ times}} h\ dx^1 \cdots dx^m = \overbrace{\underset{\psi(U)}{\int \dots \int}}^{m \text{ times}} h \cdot \det \left[\frac{\partial x^i}{\partial y^i}\right]\ dy^1 \cdots dy^m$$

Where $\det\left[\frac{\partial x^i}{\partial y^i}\right]$ is the standard Jacobian determinant used in change of coordinates. As long as we use compatible charts, this Jacobian determinant is positive and thus we don't need absolute values so this is the standard change of coordinates formula for m-fold integrals as taught in a multivariable calculus course. And with this we can finally discuss the generalized Stokes theorem!

The Generalized Stokes Theorem

Let M be an m-dimensional manifold with boundary ∂M and let ω be an $(m-1)$-form on the manifold. The Generalized Stokes Theorem states that:

$$\int_M d\omega = \int_{\partial M} \omega$$

In other words, integrating the (exterior) derivative of ω over the entire manifold is equal to integrating ω over the boundary of the manifold.

Not only is this an awesome and compact result, but it is used all over the place in calculus 1, 2, and 3. In fact, almost all of the big theorems in calculus are really just special cases of the Generalized Stokes Theorem. To show this I will list the big theorems and how they are secretly all the Generalized Stokes Theorem.

7 Examples of the Generalized Stokes Theorem

In this section we will discuss the Generalized Stokes Theorem a bit more and focus on examples of how it is used in calculus. To start off let us just put all of these formulas in a table, then we can start discussing examples of each.

The Fundamental Theorem of Calculus	$\int_a^b f'(x)dx = f(b) - f(a)$
The Fundamental Theorem of Line Integrals	$\int_C \nabla f \cdot d\vec{r} = f(B) - f(A)$
Green's Theorem	$\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial p}{\partial y}\right) dxdy = \oint_{\partial D} P\ dx + Q\ dy$
Stokes' Theorem	$\iint_S (\nabla \times \vec{f}) \cdot d\vec{A} = \oint_{\partial S} \vec{f} \cdot d\vec{\ell}$
The Divergence Theorem	$\iiint_E (\nabla \cdot \vec{f})dV = \oiint_{\partial E} (\vec{f} \cdot \hat{n})dA$

Table 1: The biggest theorems in calculus. Note that P and Q are each functions of both x and y, f is a scalar field/function, and $\vec{f}$ is a vector field. For the regions to integrate over: C is a smooth 1-dimensional curve starting at A and ending at B, D is a simply connected region in $\mathbb{R}^2$ with boundary oriented counter-clockwise, S is an oriented surface in $\mathbb{R}^3$ whose boundary is also suitably oriented (from calculus 3), and E is a solid region in $\mathbb{R}^3$ with its boundary ∂E oriented outward.

So now we can discuss examples of each of these:

7.1 The Fundamental Theorem of Calculus

In calculus 1 we learn that the Fundamental Theorem of Calculus says:

$$\int_a^b f'(x)dx = f(b) - f(a)$$

but when we get to the Generalized Stokes Theorem we realize that this is just a special case. Specifically we have a manifold $M = [a, b]$ with orientation $\eta = dx$ and boundary $\partial M = \{a, b\}$. Recognize that $T_aM = \text{span}\left\{-\frac{d}{dx}\Big|_a\right\}$ we can consider the outward facing normal on the boundary of the manifold:

$$dx\left(-\frac{d}{dx}\Bigg|_a\right) = -1 \qquad \text{and} \qquad dx\left(\frac{d}{dx}\Bigg|_b\right) = 1$$

Taking into account that a is incompatible and b is compatible we can induce an orientation that is compatible on the whole boundary:

$$\tilde{\eta} = \begin{cases} -1 & p = a \\ 1 & p = b \end{cases}$$

Now consider a 0-form $\omega = f$ such that $d\omega = f'(x)dx$. Using this and our manifold we can plug things into both sides of the Generalized Stokes Theorem:

$$\text{Left side} = \int_M d\omega = \int_a^b f'(x)dx$$

$$\text{Right side} = \int_{\partial M} \omega = \int_{\{a,b\}} f = -\int_{\{a\}} f + \int_{\{b\}} f = -f(a) + f(b)$$

This shows that the Fundamental Theorem of Calculus is a special case of the Generalized Stokes Theorem when M is a 1-dimensional manifold on $\mathbb{R}$ and ω is a 0-form.

7.2 The Fundamental Theorem of Line Integrals

Let's work through a specific example of the Fundamental Theorem of Line Integrals. Consider the curve $M : y = x^2$ with $z = 1$ parameterized using $\vec{r}(t) = \langle t, t^2, 1 \rangle$ over $0 \leq t \leq 1$. The boundary is then $\partial M = \{\vec{r}(0), \vec{r}(1)\} = \{(0,0,1), (1,1,1)\}$. Orient this using $\eta = dt$ so that on the boundary our compatible outward pointing directions are negative at 0 and positive at 1. Consider the 0-form $f = x + z$ such that $df = dx + dz$. We can apply the Generalized Stokes Theorem to this to get:

$$\int_M df = \int_{[\vec{r}(0),\vec{r}(1)]} \left[f_x(\vec{r}(t))dx(\vec{r}\,'(t)) + f_y(\vec{r}(t))dy(\vec{r}\,'(t)) + f_z(\vec{r}(t))dz(\vec{r}\,'(t))\right] dt =$$

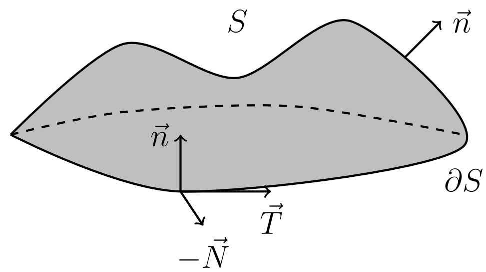
$$\int_0^1 \left[f_x(\vec{r}(t))x'(t) + f_y(\vec{r}(t))y'(t) + f_z(\vec{r}(t))z'(t)\right] dt = \int_0^1 (1+0)dt = 1$$

And the other side of the theorem gives:

$$\int_{\partial M} f = \int_{\{\vec{r}(0),\vec{r}(1)\}} f = -\int_{\{\vec{r}(0)\}} f + \int_{\{\vec{r}(1)\}} f = -f(\vec{r}(0)) + f(\vec{r}(1)) = -1 + 2 = 1$$

7.3 Stokes' Theorem (General Case)

Consider a bounded surface S with boundary ∂S in $\mathbb{R}^3$ as is shown to the right. Let $\omega = Pdx + Qdy + Rdz$ be a 1-form defined over this surface corresponding to the vector field $\vec{F} = \langle P, Q, R\rangle$. If we parameterize S using $\vec{r}(u, v)$ where $(u, v) \in D$ then calculus 3 tells us that the unit normal is $\vec{n} = \frac{\vec{r}_u \times \vec{r}_v}{|\vec{r}_u \times \vec{r}_v|}$. We can define an orientation on $\mathbb{R}^3$ and hence on S by taking $\eta(\cdot, \cdot) = dx \wedge dy \wedge dz(\vec{n}, \cdot, \cdot)$.



Checking that $\vec{r}$ is compatible with the orientation:
$\eta(\vec{n}, \vec{r}_u, \vec{r}_v) = dx \wedge dy \wedge dz(\vec{n}, \vec{r}_u, \vec{r}_v) = \vec{n} \cdot (\vec{r}_u \times \vec{r}_v)$ which we know is the volume of a parallelepiped. Note that our three vectors are not coplanar and they are consistent with the right-handed coordinate system therefore this volume is non-zero and it has positive sign therefore the orientation is compatible. At this point we have two pieces to calculate:

1. We can parameterize the boundary using $\partial S_1 : \vec{r}(t)$ for $a \le t \le b$. This lets us calculate the tangent and normal vectors as shown in the diagram, $\vec{T} = \vec{r}\,'/|\vec{r}\,'|$ and $\vec{N} = \vec{T}\,'/|\vec{T}\,'|$. Notice that $\vec{N}$ will point into the surface so we will choose $-\vec{N}$ as the outward pointing normal. We can induce an orientation on the boundary as:
$$\tilde{\eta}(\cdot) = \eta(-\vec{N}, \cdot) = dx \wedge dy \wedge dz(\vec{n}, -\vec{N}, \cdot)$$
Checking the compatibility of this orientation we get
$$\tilde{\eta}(\vec{r}\,'(t)) = dx \wedge dy \wedge dz(\vec{n}, -\vec{N}, \vec{r}\,') = \vec{n} \cdot (-\vec{N} \times \vec{r}\,') > 0$$
for similar reasons to the first orientation check we did. So now we can calculate the first side of the equation:
$$\boxed{\int_{\partial S_1} \omega = \int_a^b \omega_{\vec{r}(t)}(\vec{r}\,'(t))dt = \int_{\partial S_1} \vec{F} \cdot d\vec{r}}$$

2. Since we will need this in a moment let's calculate $d\omega$.
$$d\omega = (R_y - Q_z)dy \wedge dz - (R_x - P_z)dz \wedge dx + (Q_x - P_y)dx \wedge dy$$
Now for the other side of the equation:
$$\int_{S_1} d\omega = \iint_D (R_y - Q_z)dy\wedge dz(\vec{r}_u, \vec{r}_v) - (R_x - P_z)dz\wedge dx(\vec{r}_u, \vec{r}_v) + (Q_x - P_y)dx\wedge dy(\vec{r}_u, \vec{r}_v)$$

Since

$$(dy \wedge dz)(\vec{r}_u, \vec{r}_v) = \det \begin{bmatrix} y_u & y_v \\ z_u & z_v \end{bmatrix} = y_u z_v - z_u y_v$$

we can use this, and the corresponding simplifications from the other parts, to rewrite the equation as:

$$\int_{S_1} d\omega = \iint_D (R_y - Q_z)(y_u z_v - z_u y_v) - (R_x - P_z)(z_u x_v - x_u z_v) + (Q_x - P_y)(x_u y_v - y_u x_v)$$

Since this form is a bit hard to recognize unless you are looking for it, let's consider the curl of $\vec{F}$ to get some insights.

$$\iint_D \left(\nabla \times \vec{F}\right) \cdot \vec{n}\ d\sigma = \iint_D \left(\nabla \times \vec{F}\left(\vec{r}(u,v)\right)\right) \cdot (\vec{r}_u, \vec{r}_v)\, dA =$$

$$\iint_D \det \begin{bmatrix} R_y - Q_z & x_u & x_v \\ R_x - P_z & y_u & y_v \\ Q_x - P_y & z_u & z_v \end{bmatrix} =$$

$$\iint_D (R_y - Q_z)(y_u z_v - z_u y_v) - (R_x - P_z)(z_u x_v - x_u z_v) + (Q_x - P_y)(x_u y_v - y_u x_v)$$

This is what we got above so we can say that

$$\boxed{\int_{S_1} d\omega = \iint_D \left(\nabla \times \vec{F}\right) \cdot \vec{n}\ d\sigma}$$

Finally, we can combine both sides to evaluate the Generalized Stokes Theorem in this example:

$$\int_{S_1} d\omega = \int_{\partial S_1} \omega$$

$$\downarrow$$

$$\iint_D \left(\nabla \times \vec{F}\right) \cdot \vec{n}\ d\sigma = \int_{\partial S_1} \vec{F} \cdot d\vec{r}$$

And this is precisely the normal calculus 3 Stokes' Theorem.

7.4 Stokes' Theorem Specific Example

After looking at the general way this type of example goes, let us now consider a specific example using actual numbers:
Let our surface be $S_1 = x^2 + y^2 + z^2 = 4$ with $z \geq 0$ and oriented up. Clearly ∂S_1 gives $x^2 + y^2 = 4$ with $z = 0$ and we will assume we are oriented in the counter clockwise direction. Let $\omega = -ydx + xdy + z^2 dz$ then

$$d\omega = -dy \wedge dx + dx \wedge dy + 2zdz \wedge dz = 2dx \wedge dy$$

Parameterizing S_1 using spherical coordinates we get

$$S_1 : \vec{r}(\phi, \theta) = \langle 2\cos(\theta)\sin(\phi), 2\sin(\theta)\sin(\phi), 2\cos(\phi)\rangle$$

Where θ ranges from 0 to 2π and ϕ ranges from 0 to $\pi/2$. Using this parameterization we can determine our vectors:

$$\vec{r}_\phi = \langle 2\cos(\theta)\cos(\phi), 2\sin(\theta)\cos(\phi), -2\sin(\phi)\rangle$$

$$\vec{r}_\theta = \langle -2\sin(\theta)\sin(\phi), 2\cos(\theta)\sin(\phi), 0\rangle$$

$$\vec{r}_\phi \times \vec{r}_\theta = 4\sin(\phi)\langle \cos(\theta)\sin(\phi), \sin(\theta)\sin(\phi), \cos(\phi)\rangle$$

Note that the normal vector $\vec{n} = \langle \cos(\theta)\sin(\phi), \sin(\theta)\sin(\phi), \cos(\phi)\rangle$ does point outwards. If we now consider an orientation for the surface we get

$$\eta(\cdot,\cdot) = dx \wedge dy \wedge dz(\vec{n}, \cdot, \cdot) = \cos(\theta)\sin(\phi) dy \wedge dz + \sin(\theta)\sin(\phi) dz \wedge dx + \cos(\phi) dx \wedge dy$$

Which simplifies to

$$\frac{1}{2}\left(x dy \wedge dz + y dz \wedge dx + z dx \wedge dy\right)$$

Also note that $\eta(\vec{r}_\phi, \vec{r}_\theta) > 0$ over the domain. Now we can finally get to calculating both sides of the equation:

1. Starting with the integral over the entire manifold:

$$\int_{S_1} d\omega = \int_{S_1} 2dx \wedge dy = \int_0^{2\pi}\int_0^{\pi/2} 2dx \wedge dy(\vec{r}_\phi, \vec{r}_\theta) d\phi d\theta =$$

$$\int_0^{2\pi}\int_0^{\pi/2} \det\begin{bmatrix} 2\cos(\theta)\cos(\phi) & -2\sin(\theta)\sin(\phi) \\ 2\sin(\theta)\cos(\phi) & 2\cos(\theta)\sin(\phi)\end{bmatrix} d\phi d\theta =$$

$$\int_0^{2\pi}\int_0^{\pi/2} 2\cdot(4\sin(\phi)\cos(\phi)) d\phi d\theta = 8\pi \sin^2(\phi)\Big|_0^{\pi/2} = 8\pi$$

2. Now for the other side. First parameterize the boundary using $\partial S_1 : \vec{r}(t) = \langle 2\cos(t), 2\sin(t), 0\rangle$ then $\vec{r}\,'(t) = \langle -2\sin(t), 2\cos(t), 0\rangle$. Geometrically we can see that $-\hat{k}$ is the outward pointing normal so the induced orientation would be

$$\tilde{\eta}(\cdot) = \eta(-\hat{k}, \cdot) = \frac{1}{2} x dy \wedge dz(-\hat{k}, \cdot) + \frac{1}{2} y dz \wedge dx(-\hat{k}, \cdot) = \frac{x}{2} dy - \frac{y}{2} dx$$

Checking compatibility: $\tilde{\eta}(\vec{r}\,'(t)) = \frac{2\cos(t)}{2}\cos(t) - \frac{2\sin(t)}{2}\sin(t) = 2 > 0$. Now we can calculate the other side:

$$\int_{\partial S_1} \omega = \int_0^{2\pi} \omega_{\vec{r}(t)}(\vec{r}\,'(t)) = \int_0^{2\pi} (-2\sin(t)\cdot(-2\sin(t)) + 2\cos(t)\cdot(\cos(t))) dt =$$

$$\int_0^{2\pi} 4dt = 8\pi$$

And thus we have verified Stokes' Theorem for this example.

7.5 The Divergence Theorem (General Case)

Take our manifold to be some region E in local coordinates (x, y, z) with boundary ∂E. Let $\omega = Pdy \wedge dz + Qdz \wedge dx + Rdx \wedge dy$ be a 2-form using $\vec{F} = \langle P, Q, R\rangle$. We can orient E, and actually all of $\mathbb{R}^3$ using $\eta = dx \wedge dy \wedge dz$. Note that is compatible because $\eta\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) = \det(I^3) = 1 > 0$. Calculating $d\omega$ we get

$$d\omega = (P_x + Q_y + R_z)dx \wedge dy \wedge dz$$

Computing the integral of $d\omega$ over the entire region we get:

$$\int_E d\omega = \iiint_E (P_x + Q_y + R_z)dx \wedge dy \wedge dz\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) dV =$$

$$\iiint_E (P_x + Q_y + R_z)dV = \iiint_E \left(\nabla \cdot \vec{F}\right) dV$$

For the other side of the theorem we will first need to parameterize the boundary, $\partial E : \vec{r}(u, v)$ where $(u, v) \in D$ such that $\vec{n} = \frac{\vec{r}_u \times \vec{r}_v}{|\vec{r}_u \times \vec{r}_v|}$. Inducing an orientation on the boundary we get $\tilde{\eta}(\cdot, \cdot) = \eta(\vec{n}, \cdot, \cdot) = dx \wedge dy \wedge dz(\vec{n}, \cdot, \cdot)$. Checking the compatibility we find $\tilde{\eta}(\vec{r}_u, \vec{r}_v) = dx \wedge dy \wedge dz(\vec{n}, \vec{r}_u, \vec{r}_v) = \vec{n} \cdot (\vec{r}_u, \vec{r}_v)$ which is the volume of a parallelepiped where the three vectors are not coplanar and since the coordinate system is right handed the value must be positive. Finally we can calculate the other side of the theorem:

$$\int_{\partial E} \omega = \iint_D \omega_{\vec{r}(u,v)}(\vec{r}_u, \vec{r}_v)dA =$$

$$\iint_D P(\vec{r}(u, v))\det\begin{bmatrix} y_u & y_v \\ z_u & z_v \end{bmatrix} + Q(\vec{r}(u, v))\det\begin{bmatrix} z_u & z_v \\ x_u & x_v \end{bmatrix} + R(\vec{r}(u, v))\det\begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} =$$

$$\iint_D \det\begin{bmatrix} P & x_u & x_v \\ Q & y_u & y_v \\ R & z_u & z_v \end{bmatrix} dA = \iint_D \vec{F} \cdot (\vec{r}_u \times \vec{r}_v)dA = \iint_D \vec{F} \cdot \vec{n}\ d\sigma$$

Combining both sides of the equation we get the divergence theorem:

$$\iiint_E \left(\nabla \cdot \vec{F}\right) dV = \iint_D \vec{F} \cdot \vec{n}\ d\sigma$$

References

[C] Susan Colley. *Vector Calculus*. Prentice Hall, 1998.

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

[J] Nathan Jacobson. *Lectures in Abstract Algebra: Volume III – Theory of Fields and Galois Theory*. D. Van Nostrand Company, Inc., Princeton, NJ, 1964.

[MAA] Victor J. Katz. *The History of Stokes' Theorem*. Mathematical Association of America. Published in Mathematics Magazine, Vol. 52, No. 3 (May, 1979) pp. 146-156

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

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

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

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

[W] Wikipedia: Manifold, https://en.wikipedia.org/wiki/Manifold