TRINITY COLLEGE DUBLIN JUNIOR SOPHISTER THEORETICAL PHYSICS

Notes on MAU33206 - Calculus on Manifolds

Matthew Dunne

- **Moderators**: Florian Naef (Lecturer) and Rodrigo Navarro (Tutor)
- **Grade Weighting**: 10% CA (Weekly Homeworks) and 90% Exam.
- **Textbook References**: Analysis on Manifolds - Munkres

CONTENTS

We begin this course with a review of the ingredients needed to make contact with the standard definition of a manifold. These ingredients include concepts from linear algebra and analysis in several real variables.

1 Review of the Topology of $\mathbb{R}^n$

A quick run through of the definitions and theorems seen in the study of $\mathbb{R}^n$.

Definition (Metric Space): Let X be a set in $\mathbb{R}^n$. A **metric** on X is a function $d : X \to \mathbb{R}$ such that:

- $d(x, y) = d(y, x)$
- $d(x, y) \geq 0$ with equality iff $x = y$
- $d(x, z) \leq d(x, y) + d(y, z)$ (Triangle Inequality)

The set X together with the metric d forms a **metric space** denoted (X, d)

Definition (Open and Closed Subsets in Metric Spaces): Let (X, d) be a metric space. Then $U \subseteq X$ is **open** if for every point $p \in U$ there exists an $\varepsilon > 0$ such that $B_\varepsilon(p) \subseteq U$. The set U is **closed** if $X - U$ is open.

Proposition 1.1: Let (X, d) be a metric space with $U \subseteq Y \subseteq X$. Then U is **open/closed** in $(Y, d|_{Y\times Y})$ if and only if there exists a set $V \subseteq X$ **open/closed** such that $U = V \cap Y$.

Definition (Continuity of Mappings Between Metric Spaces): Let X and Y be metric spaces and $f : X \to Y$ be a mapping from X to Y. Then f is **continuous** if $f^{-1}(U)$ is open whenever $U \subseteq Y$ is open. If in addition, the inverse function f^{-1} is continuous and f itself is a bijection, then we call it a **homomorphism**.

Definition (Compactness): A set $X \subseteq \mathbb{R}^n$ is **compact** if every open cover of X has a finite subcover. i.e. For all collections $\{U_\alpha\}_{\alpha\in I}$ where $U_\alpha \subseteq X$ such that $\bigcup_{\alpha\in I} U_\alpha = X$, there exists $\alpha_1, \alpha_2, \ldots, \alpha_k \in I$ such that $X = U_{\alpha_1} \cup U_{\alpha_2} \cup \cdots \cup U_{\alpha_k}$.

Theorem (Heine-Borel): A set $X \subseteq \mathbb{R}^n$ is compact if and only if it is both bounded and closed.

Definition (Differentiability): Let $U \subseteq \mathbb{R}^n$ be open and $f : U \to V \subseteq \mathbb{R}^m$ be a function from $\mathbb{R}^n$ to $\mathbb{R}^m$. Then f is said to be **differentiable** at $p \in U$ with derivative $Df(p)$ if

$$\lim_{x\to p} \frac{f(x) - f(p) - Df(p)(x - p)}{||x - p||} = 0$$

Differentiability Classes:

- C^0: All continuous functions.
- C^1: Functions whose derivative both exists everywhere and is of class C^0 i.e. $Df : U \to Mat_{m\times n} \cong \mathbb{R}^{mn}$.

- C^r: The set of all differentiable functions whose derivative is in C^{r-1} i.e. all partials of all components are r times continuously differentiable.
- C^∞ (smooth): C^k for all $k > 0$.
- C^ω (analytic): Functions that are smooth and whose Taylor series expansion around any point in its domain converges to the function in some neighborhood of the point.

Proposition 1.2 (Smoothness in terms of Derivatives): For a set $U \subseteq \mathbb{R}^n$, a function $f : U \to \mathbb{R}^m$ mapping x to $(f_1(x), \dots, f_m(x))$ is C^r if and only if the derivatives

$$\frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_k}} f_j : U \to \mathbb{R}.$$

exist and are continuous for all $k \in \{1, \dots, r\}$, $i_1, \dots, i_k \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$. In this case Df is the $m \times n$ matrix whose elements are given by $\left(\frac{\partial f_j}{\partial x_{i_k}}\right)$.

Theorem (Multidimensional Chain Rule): If the functions $g : U \to V$ and $f : V \to W$ are differentiable or C^r then so is $f \circ g$ and

$$D(f \circ g)(x) = Df(g(x))Dg(x).$$

Theorem (Implicit Functions Theorem): Let $f : V \to \mathbb{R}^n$ map the open set $V \subseteq \mathbb{R}^n$ into $\mathbb{R}^n$. Let f be C^r where $(1 \le r \le \infty)$ and $p \in V$. Suppose $Df(p)$ is non-singular (an invertible $n \times n$ matrix), then there exists an open neighbourhood of p denoted by $U \subseteq V$ such that $f|_U : U \to f(U)$ is a C^r-diffeomorphism.

2 Manifolds in $\mathbb{R}^n$

The informal intuitive way to think about what a manifold is, is a subset of Euclidean space which is smooth, without corners or creases and does not intersect itself. We shall now make some rigorous definitions of the concept for standard manifolds in $\mathbb{R}^n$.

Definition (d-Manifold Without Boundary in $\mathbb{R}^n$): Let $d > 0$. A subset $M \subseteq \mathbb{R}^n$ is called a d-manifold without boundary in $\mathbb{R}^n$ of class C^r if for all $p \in M$, there exists an open subset $V \subseteq M$ of which p is an element, a set U that is open in $\mathbb{R}^d$, and a continuous map $\alpha : U \to V$ carrying U onto V in a one-to-one fashion such that:

- α is of class C^r
- $\alpha^{-1} : V \to U$ is continuous.
- $D\alpha(x)$ is rank-d for all $x \in U$.

The map α is called a **coordinate patch** on M about p. If such an alpha exists and works for all points p, then M is called a parameterised manifold.

Example 2.1 (d=1): Consider the subset $M = \{(t^3, t^2) \mid t \in \mathbb{R}\}$ of $\mathbb{R}^2$. The mapping from $\mathbb{R}$ ($d = 1$) for which M is the image set is defined as $\alpha : \mathbb{R} \to M$ such that $t \mapsto (t^3, t^2)$. This one-to-one function is of class C^∞ with a continuous inverse and is therefore a homomorphism. Its derivative, $D\alpha(t) = (3t^2, 2t)$ vanishes at $(0,0) = \mathbf{0}$ meaning that M must have a cusp at the origin. Combined with the fact that M is the exact image set of α we can conclude that no coordinate patch exists such that its derivative preserves a rank of 1 at the origin.

Remark (Proving Coordinate Patch): Since the condition that $D\alpha(x)$ must be rank-d for all $x \in U$ is the most important for making definite proofs about whether a function is a coordinate patch or not. It is useful to walk through some of the low dimensional cases which we will mostly work with as the condition can be thought of slightly differently for each $d = 1, 2, 3, \ldots$.

- $(d = 1) \iff$ For any point $x \in U$, $D\alpha(x)$ is rank-1, $\iff$, $D\alpha(x) \neq \mathbf{0}$.
- $(d = 2) \iff$ For any point $a \in V$, $D\alpha(a)$ is rank-2, $\iff$ the columns $\frac{\partial \alpha}{\partial x_1}$ and $\frac{\partial \alpha}{\partial x_2}$ of $D\alpha$ are linearly independent at a. In which case the derivatives also form a tangent plane to the surface of M at the point a.

Definition (C^r Function Class): Let $S \subseteq \mathbb{R}^m$ and f be a function mapping S to $\mathbb{R}^n$. We say f is C^r if it may be extended to a function $g : U \to \mathbb{R}^n$ that is C^r on an open set $U \subseteq \mathbb{R}^m$ containing S. This also leads to **compositions** of C^r functions being C^r.

Definition (Smooth Functions): Let $M \subseteq \mathbb{R}^n$ and $N \subseteq \mathbb{R}^m$ be subsets. A function $f : M \to N$ is **smooth** if there exists a set $M \subseteq V \subseteq_{open} \mathbb{R}^n$ and $\tilde{f} : V \to \mathbb{R}^m$ smooth such that $\tilde{f}$ is an extension of f, i.e. $\tilde{f}|_M = f$.

This leads to the following two conditions being equivalent:

- $f : M \to N$ is smooth and has a smooth inverse f^{-1}.
- There exists an extension of f, namely $\tilde{f}$ which is smooth and whose inverse $\tilde{f}^{-1}$ is smooth and an extension of f^{-1} such that $\tilde{f}^{-1} \circ \tilde{f}|_M = Id_M$ and $\tilde{f} \circ \tilde{f}^{-1}|_N = Id_N$.

Definition (Diffeomorphism): If $U \subseteq \mathbb{R}^m$ and $V \subseteq \mathbb{R}^n$ are both open and a function $f : U \to V$ is smooth such that $f^{-1} : V \to U$ exists and is also smooth, then f is a **diffeomorphism**.

Remark (Restrictions of Diffeomorphisms): Let $M \subseteq \mathbb{R}^n$ and $N \subseteq \mathbb{R}^m$ be subsets and let $f : M \to N$ be a diffeomorphism. Then the restriction $f|_A : A \to f(A)$ where $A \subseteq M$ is **also a diffeomorphism**.

From the definition of smooth functions and diffeomorphisms it follows that **compositions of smooth functions are smooth and compositions of diffeomorphisms are diffeomorphisms**.

Theorem 2.2: Let $M \subseteq \mathbb{R}^n$ be a set, $d > 0$, and $p \in M$. The following statements are equivalent.

- There exists a set $V \ni p$ open in M and U open in $\mathbb{R}^d$ and a smooth (C^∞) homomorphism $\alpha : U \to V$ such that $D\alpha(x)$ is rank-d at all points $x \in U$.
- There exists a set $V \ni p$ open in $\mathbb{R}^n$ and U open in $\mathbb{R}^n$ with a mapping $\beta : U \to V$ that is a diffeomorphism such that $\beta(U \cap (\mathbb{R}^d \times \{0\})) = V \cap M$.

Proof: "$\Longrightarrow$": Without loss of generality, we assume that $\alpha(0) = p$, then since $D\alpha(x)$ has rank-d for any $x \in U$, there must exist some invertible matrix $B \in Mat_{n\times(n-d)}$ such that $[D\alpha(0), B]$ is non-degenerate (all columns linearly independent). Now we cleverly choose to define $\beta : U \times \mathbb{R}^{n-d} \to \mathbb{R}^n$ as $(x, y) \mapsto \alpha(x) + By$ which clearly implies $D\beta(0,0) = [D\alpha(0), B]$.

Now by the inverse function theorem, we demand there exist sets $\tilde{U} \ni 0$ and $\tilde{V}$ open in $\mathbb{R}^n$ such that $\beta : \tilde{U} \to \tilde{V}$ is a diffeomorphism. Then $\beta(\tilde{U} \cap \mathbb{R}^d \times \{0\}) = \alpha(\tilde{U} \cap \mathbb{R}^d \times \{0\})$ (an open neighbourhood of p) where both of these image sets are open in M. Now letting $(\tilde{U} \cap \mathbb{R}^d \times \{0\}) = M \cap \tilde{\tilde{V}}$ for some $\tilde{\tilde{V}}$ open in $\mathbb{R}^n$, we then make the replacements $\tilde{V} \longleftrightarrow (\tilde{V} \cap \tilde{\tilde{V}})$ and $\tilde{U} \longleftrightarrow \beta^{-1}(\tilde{V} \cap \tilde{\tilde{V}})$ which gives that $\beta(\tilde{U} \cap \mathbb{R}^d \times \{0\}) = \tilde{V} \cap M$.

"$\Longleftarrow$": Let $\tilde{U} = U \cap \mathbb{R}^d \times \{0\}$ and $\beta : U \to V$ be a diffeomorphism (U and V open in $\mathbb{R}^n$). Let α be the restriction of β to the set $\tilde{U}$, $\alpha = \beta|_{\tilde{U}} : \tilde{U} \to \beta(U) \cap M$. Restricting a diffeomorphism gives a homomorphism so α is a homomorphism onto $\beta(U) \cap M$ open. Therefore $\alpha(x_1, \ldots, x_d) = \beta(x_1, \ldots, x_d, 0, \ldots, 0)$ and

$$D\alpha = D\beta \cdot \begin{pmatrix} \mathbb{I}_{d\times d} \\ \mathcal{O} \end{pmatrix}$$

where the matrix is the $n \times d$ matrix whose upper $d \times d$ block is the identity and 0 everywhere else. So clearly $D\alpha$ has rank-d. □

This powerful theorem shows that when we say that a d-manifold without boundary must have a smooth, homomorphic coordinate patch at each point which maps into the neighbourhood of the point with a rank d derivative, we can really just say that the coordinate patch must be a diffeomorphism and call it a day.

Corollary (Alternate Definition of a d-Manifold Without Boundary): Let M be a subspace of $\mathbb{R}^n$. Then M is a *d-Manifold without boundary* iff for every point $p \in M$ there exists an open neighbourhood $p \in V \subseteq M$, an open neighbourhood of d-dimensional space $U \subseteq \mathbb{R}^d$ and a diffeomorphism $\alpha : U \to V$.

Proof: " $\implies$ ": If $\beta : U \to V$ is a diffeomorphism, then the restriction $\beta|_{U \cap \mathbb{R}^d \times \{0\}}$ is a diffeomorphism onto $V \cap M$ which contains p.

" $\impliedby$ ": Letting $\alpha : U \to V$ be a diffeomorphism. Then α is a smooth homomorphism by definition. Also this implies that $\alpha^{-1} : V \to U$ is smooth and therefore has a smooth extension for some $\tilde{V}$ containing V, $\tilde{\alpha}^{-1} : \tilde{V} \to \mathbb{R}^d$ such that $\tilde{\alpha}^{-1}|_V = \alpha^{-1}$. It follows immediately that $\tilde{\alpha}^{-1} \circ \alpha = \alpha^{-1} \circ \alpha = Id_U$. Then by the chain rule the derivative is $D\tilde{\alpha}^{-1}(\alpha(x)) \cdot D\alpha = \mathbb{I}_{d \times d}$ which means $D\alpha$ is rank-d. $\square$

Now we will move to expanding out the more local conditions of a manifold to make them more easy to work with. The following lemma will guide us in linking together all coordinate patches for different points of a manifold.

Lemma 2.3 (Locally Smooth $\implies$ Smooth: Let $M \subseteq \mathbb{R}^m$, and $f : M \to \mathbb{R}^n$. If for each $x \in M$, there is a neighbourhood U_x of x and a function $g_x : U_x \to \mathbb{R}^n$ of class C^r that agrees with f on $U_x \cap M$, then f is C^r on M.

Definition (Support): If $\phi : \mathbb{R}^n \to \mathbb{R}$, then the **support** of ϕ is the closure of the set $\{x \mid \phi(x) \neq 0\}$. In other words any point not in the support has a neighbourhood on which ϕ vanishes.

Theorem 2.4 (Existence of Partition of Unity): Let $\mathcal{A}$ be a collection of open sets in $\mathbb{R}^n$ and let their union be $V = \bigcup_{\alpha \in \mathcal{A}} U_\alpha$. There exists a sequence $\phi_1, \phi_2, \ldots$ of continuous functions $\phi_i : \mathbb{R}^n \to \mathbb{R}$ such that:

1. $\phi_i(x) \geq 0$ for all x.
2. The set $S_i = Support(\phi_i)$ is contained in V.
3. Each point of V has a neighbourhood that intersects only finitely many S_i sets.
4. $\sum_{i=1}^{\infty} \phi_i(x) = 1$ for all x.
5. ϕ_i are smooth functions.
6. S_i are compact (closed and bounded).
7. For each i, the set S_i is contained in $\mathcal{A}$.

A collection of functions $\{\phi_i\}$ satisfying conditions (1) - (4) is called a **partition of unity**. If it satisfies (5) then its said to be C^∞. If it satisfies (6) its said to have **compact supports**. If it satisfies (7) it is **dominated by the collection $\mathcal{A}$**.

Just note that the sum in condition (3) is locally finite due to (2). The proof of this theorem will follow shortly.

Lemma 2.5: Given a collection of sets with union $V = \bigcup_{\alpha \in \mathcal{A}} U_\alpha$ as above in (2.4). There exists points $\{p_1, p_2, \dots\} \in \mathbb{R}^n$ and strictly positive real numbers $\{\varepsilon_1, \varepsilon_2, \dots\} \in \mathbb{R}$ such that:

1. $\bigcup_{i=1}^{\infty} B_{\varepsilon_i}(p_i) = V$.
2. Each open ball of double the radius $B_{2\varepsilon_i}(p_i)$ is contained in one of the sets U_α.
3. Each point $p \in V$ has a neighbourhood intersecting finitely many open balls $B_{2\varepsilon_i}(p_i)$.

Sub-lemma: One can always find a nested sequence of sets $K_1 \subseteq K_2 \subseteq \dots \subseteq V$ such that K_i are **compact**, and the union $\bigcup_i K_i = V$.

Proof of Sub-lemma: Let K_i be the set, $\overline{B_i(0)} - \bigcup_{x \in V^c} B_{2^{-i}}(x)$. As we let i run through the positive integers we see that we indeed obtain a nested sequence of compact sets centred on the point 0 which was chosen arbitrarily. But we have eliminated points just within very small neighbourhoods of V^c and so when we take the union of all the K_i's we get back V. $\square$

Proof of Lemma 2.5: Set $K_p = K_{-1} = \phi$. Then we cover the sets $K_i - K_{i-1}^c$ with $B_{\varepsilon_x}(x)$ such that

1. $B_{2\varepsilon_x}(x) \subset K_{i-2}^c$.
2. $B_{2\varepsilon_x}$ is contained in some U_x.

By compactness we have finitely many x_j^i, ε_j^i, such that $B_{\varepsilon_j^i}(x_j^i)$ covers $K_i - K_{i-1}^c$. Now we claim that the values $\{x_j^i, \varepsilon_j^i\}$ that we have constructed satisfy the properties in the lemma.

1. $\bigcup_i K_i - K_{i-1}^c = V$
2. Proven by construction
3. $B_{2\varepsilon_j^i}(x_j^i)$ does not intersect $K_{i-2} \supset K_{i-1} \supset \cdots$ and therefore K_i intersects finitely many balls $B_{2\varepsilon_j^i}(x_j^i)$. Now for any $p \in A$ there exists an integer i such that $p \in K_i \subset K_{i+1}$ and K_{i+1} is the open neighbourhood intersecting only finitely many $B_\varepsilon(x)$'s. $\square$

Lemma 2.6: Let $p \in \mathbb{R}^n$, $\varepsilon > 0$. Then, there exists $\varphi : \mathbb{R}^n \to [0, 1]$ such that:

1. φ is smooth.

2. $\varphi \subset B_{2\varepsilon}(p)$.

3. $\varphi > 0$ on $B_{2\varepsilon}(p)$.

Proof: WLOG we can let $p = 0$, $\varepsilon = \frac{2}{3}$ and then defining φ as

$$\varphi(x) = \begin{cases} e^{\frac{1}{1-\|x\|^2}} & \text{if } \|x\| < 1 \\ 0 & \text{if } otherwise \end{cases}.$$

The results follow.

Proof of Theorem 2.4: Find p_i, ε_i as in Lemma 2.5, and find $\tilde{\varphi}_i$ as in Lemma 2.6. With these selections we can verify (1) - (5) but not (3) and then modify $\tilde{\varphi}_i$.

1. By Lemma 2.6. we have Support $\tilde{\varphi}_i \subset B_{2\varepsilon_i}(p_i) \subset U_\alpha$ for some α.

2. The set S_i = Support $\tilde{\varphi}_i$ is contained within $B_{2\varepsilon_i}(p_i)$ and $B_{2\varepsilon_i}(p_i)$ have the local finiteness property by Lemma 2.5.

4. By Lemma 2.6, S_i being closed and contained within bounded sets means that S_i are compact.

5. By Lemma 2.6. $\tilde{\varphi}_i$ are smooth.

Now to prove (3) we consider the claim $\psi(x) = \sum_{i=1}^{\infty} \tilde{\varphi}_i(x) > 0$ for $x \in V$. First let $x \in V$, then by Lemma 2.5, $x \in B_{\varepsilon_i}(p_i)$ for some i and hence $\psi(x) \geq \sum_{i=1}^{\infty} \tilde{\varphi}_i(x) > 0$. Now note that we can set $\varphi_i(x) = \frac{\tilde{\varphi}_i(x)}{\psi(x)}$ while still satisfying the other conditions. It follows that

$$\sum \varphi_i(x) = \sum_{i=1}^{\infty} \frac{\tilde{\varphi}_i(x)}{\psi(x)} = \frac{1}{\psi(x)} \sum_{i=1}^{\infty} \tilde{\varphi}_i(x) = 1.$$

With some of the messiest proofs I've ever seen now being completed, we can reap the rewards with the following groundbreaking proposition which can help us with coordinate patches.

Proposition 2.7 (Locally Smooth $\implies$ Smooth): Let $M \subseteq \mathbb{R}^n$ be a subset and $f : M \to \mathbb{R}^m$ be a function. Suppose f is locally smooth, meaning for every point $p \in M$ there is an open neighbourhood V containing p on which f is smooth. Then f is smooth.

Proof: For each $x \in M$ we find a set $x \in V_x \subset_{open} \mathbb{R}^n$ and a function $\tilde{f}_x : V_x \to \mathbb{R}^m$ smooth such that $\tilde{f}_x|_{V_x \cap M} = f|_{V_x \cap M}$. Let $\{\phi_i\}$ be a partition of unity subordinate to $\{V_x\}_{x \in M}$. For each i there exists a V_{x_i} such that Support $\phi_i \subset V_{x_i}$. Now define the functions

$$h_i(x) = \begin{cases} \phi_i(x)\tilde{f}_{x_i}(x) & \text{if } x \in V_{x_i} \\ 0 & \text{if } otherwise \end{cases}.$$

Then since the function has a factor being in the partition of unity, the support being contained within the domain ensures that $h_i(x)$ will fall off smoothly before reaching the boundary of the neighbourhood V_x around x and combined with the fact that $\tilde{f}(x)$ is smooth, we find that $h_i : \bigcup_{x \in M} V_x \to \mathbb{R}^m$ is smooth. Now define $\tilde{f}(x) = \sum_{i=1}^{\infty} h_i(x)$. For each $x \in \bigcup V_x$, there exists an open neighbourhood V intersecting finitely many S_i's, namely $S_1, \dots, S_N$. Thus $\tilde{f}|_V = \sum_{i=1}^{N} h_i(x)$ and hence $\tilde{f}|_V$ is smooth. Moreover if $x \in V \cap M$, substituting the definition of $h_i(x)$ gives

$$\tilde{f}(x) = \left(\sum_{i=1}^{N} h_i(x) \right) f(x).$$

Finally since the collection forming a partition of unity sums to 1 we obtain $\tilde{f}|_M = f$ and then f is smooth by definition. $\square$

3 Manifolds With Boundary

In this section we will discuss how manifolds with a boundary are more intuitive and there are many more examples that we can work with than we had in the previous section.

Definition (Upper-Half Space): The upper half-space denoted $\mathbb{H}^d \subset \mathbb{R}^d$ is defined as $\mathbb{H}^d = \mathbb{R}^{d-1} \times \mathbb{R}_{\geq 0}$. In other words it is all points of a space for which the last component is non-negative.

The **boundary of the upper half-space** is denoted $\partial\mathbb{H}^d = \mathbb{R}^{d-1} \times \{0\}$.

The **interior of, or open version of** the upper half-space is denoted $\mathring{\mathbb{H}}^d = \mathbb{H}^d - \partial\mathbb{H}^d = \{x \in \mathbb{R}^d \mid x_d > 0\}$.

Now we will see that a manifold without boundary is no more than a special case of a manifold where the domains of its coordinate patches are open in $\mathbb{R}^d$.

Definition (d-Manifold with Boundary): A subset $M \subseteq \mathbb{R}^n$ is a d-manifold with boundary if it is locally diffeomorphic to open subsets of $\mathbb{H}^d$. That is, for any $p \in M$, there exists an open subset $V \subseteq M$ containing p, an open subset $U \subseteq \mathbb{H}^d$ and a diffeomorphism $\alpha : U \to V$.

Lemma 3.1: Let $U \subseteq \mathbb{H}^d$ be open in $\mathbb{H}^d$ but not open in $\mathbb{R}^d$. Let $\alpha : U \to \mathbb{R}^n$ be a smooth function with an extension $\tilde{\alpha} : \tilde{U} \to \mathbb{R}^n$ to the set $\tilde{U}$ open in $\mathbb{R}^d$. Then for $x \in U$, $D\tilde{\alpha}(x)$ is independent of $\tilde{\alpha}$ so we will denote it by $D\alpha(x)$ without ambiguity.

Proof: Given that $\tilde{\alpha}$ is an extension of α, we can say that the two functions agree on the set $U \subseteq \mathbb{H}^d$. Now to compute any partial derivative $\frac{\partial \tilde{\alpha}_j}{\partial x_i} = (D\tilde{\alpha}(x))_{ij}$ we form the difference quotient

$$\lim_{h \to 0} \frac{\tilde{\alpha}(x + he_i) - \tilde{\alpha}(x)}{h}$$

where e_j is the $j'th$ Euclidean standard basis vector. It suffices to take the limit as h goes to zero through positive values, in which case, if $x \in \mathbb{H}^d$ (even at the boundary) then so is $x + he_j$ (see figure below).

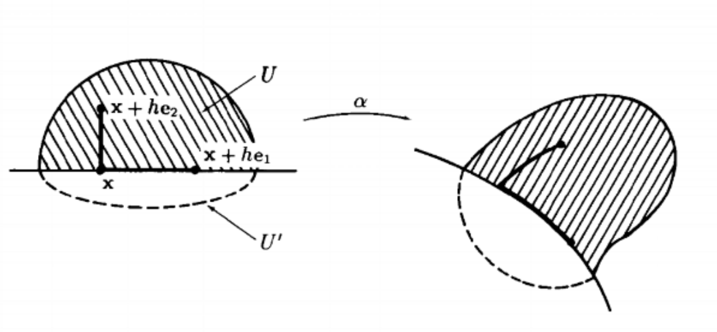

Figure 3.1: Point in the difference quotient being contained in the same set.

Note that the reason we approach 0 from positive values is to cover the case where $i = d$ in which case $x + he_d$ is not in $\mathbb{H}^d$ if h is negative. Then we have that $\tilde{\alpha}(x) = \alpha(x)$ and $\tilde{\alpha}(x + he_i) = \alpha(x + he_i)$ and therefore the difference quotient is independent of the extension $\tilde{\alpha}$ for all points. In particular $D\tilde{\alpha}(x)$ depends only on α for $x \in \mathbb{H}^d$. $\square$

Proposition 3.2: The condition for α given above in the definition of a d-manifold with boundary is equivalent to α being a smooth homeomorphism with $D\alpha$ rank-d for all $x \in U \subset \mathbb{H}^d$.

Proof: " $\Longrightarrow$ " Suppose α is a diffeomorphism between open sets $U \subseteq \mathbb{H}^d$ to $V \subseteq \mathbb{R}^n$, with inverse $\alpha^{-1} : V \to U$ then by the definition of a diffeomorphism these are smooth functions with smooth extensions . Let $\tilde{\alpha}$ and $\tilde{\alpha}^{-1}$ be smooth extensions of α and its inverse. Then $(\tilde{\alpha}^{-1} \circ \tilde{\alpha})(x) = x$ for all $x \in U \subseteq \mathbb{H}^d$. Thus we can write their composition as $(\tilde{\alpha}^{-1} \circ \tilde{\alpha})|_U(x) = Id|_U(x) : U \to U$. By the chain rule we take the derivative for $x \in \mathbb{H}^d$

$$D\tilde{\alpha}^{-1}(\tilde{\alpha}(x)) \circ D\tilde{\alpha}(x) = DId(x) = \mathbb{I}_{d \times d}$$

which implies that $D\tilde{\alpha}$ is necessarily rank-d. Then by Lemma 3.1, the restriction of this derivative to points in U is also rank$-d$ $\Longrightarrow$ $D\alpha$ rank$-d$ for all $x \in U$.

" $\Longleftarrow$ " : We use exactly the same approach as in the forward proof of Theorem 2.2. Suppose α is smooth homeomorphism with rank$-d$ derivative $D\alpha(x)$ for all $x \in U \subset \mathbb{H}^d$.

Then since $D\alpha(x)$ has rank-d for any $x \in U$, there must exist some invertible matrix $B \in Mat_{n\times(n-d)}$ such that $[D\alpha(0), B]$ is non-degenerate (all columns linearly independent i.e. rank$-n$). Now we cleverly choose to define $\beta : U \times \mathbb{H}^{n-d} \to \mathbb{R}^n$ as $(x, y) \mapsto \alpha(x) + By$ which clearly implies $D\beta(x, 0) = [D\alpha(x), B]$. So we have constructed a function with a rank$-n$ derivative for all $(x, y) \in U \times \mathbb{H}^{n-d}$.

Now by the inverse function theorem, we demand there exists two sets $\tilde{U} \subseteq U \times \mathbb{H}^{n-d}$ and $\tilde{V}$ open in $\mathbb{R}^n$ such that $\beta : \tilde{U} \to \tilde{V}$ is a diffeomorphism. Then since α is given by the restriction $\beta|_{U\times\{0\}} : U \times \{0\} \to \mathbb{R}^n$ we see that α is a diffeomorphism and given any $p \in M$, α is diffeomorphic to a neighbourhood $V \cap M \subset_{open} \mathbb{R}^n$ containing p. $\square$

Proposition 3.3 (Transition Functions are Diffeomorphisms): Let $M \subseteq \mathbb{R}^n$ be a d-manifold with boundary covered by two coordinate patches $\alpha_i : U_i \to V_i \quad i = 1, 2$. Then the **transition function** $\alpha_2^{-1} \circ \alpha_1 : \alpha_1^{-1}(V_1 \cap V_2) \to \alpha_2^{-1}(V_1 \cap V_2)$ is a diffeomorphism.

This result is important as it ensures that our coordinate patches *overlap differentiably.*

Proof: Since α_1 is a diffeomorphism, the restriction $\alpha_1|_{\alpha_1^{-1}(V_1\cap V_2)} : \alpha_1^{-1}(V_1 \cap V_2) \to V_1 \cap V_2$ is also a diffeomorphism and the same applies for α_2. Then $\alpha_2^{-1} \circ \alpha_1$ is a diffeomorphism as a composition of diffeomorphisms. $\square$

Definition (Interior and Boundary of a Manifold): Let $M \subseteq \mathbb{R}^n$ be a manifold. We call $p \in M$ an **interior point** if there exists a coordinate patch α from $U \underset{open}{\subseteq} \mathbb{H}^d$ to $V \underset{open}{\subseteq} M$ containing p such that $\alpha^{-1}(p) \in \mathring{\mathbb{H}}^d$. We call $p \in M$ a **boundary point** if $\alpha^{-1}(p) \in \partial\mathbb{H}^d$. Then we can define the interior and boundary of a manifold as

- $\mathring{M} = \{x \in M \mid \text{x an interior point}\}$
- $\partial M = \{x \in M \mid \text{x a boundary point}\}$

Note that these are not generally the interior and boundary of M in $\mathbb{R}^n$.

We now know that both the interior and the boundary are subsets of a manifold M but we are yet to discover whether their union actually constitutes the whole manifold so we must address this formally.

Lemma 3.4: A manifold with boundary M is equal to the disjoint union $\mathring{M} \cup \partial M$.

Proof: It is already clear that $\mathring{M} \cup \partial M \subseteq M$ since all boundary points and interior points are points in M and conversely every point in M is either a boundary point or an interior point. What we do not yet know is whether their union is disjoint and thus we are left to show that $\mathring{M} \cap \partial M = \varnothing$. Suppose there exists a point $p \in \mathring{M} \cap \partial M$, then there exists coordinate patches $\alpha_1 : U_1 \to V_1$ and $\alpha_2 : U_2 \to V_2$ along with points $x_1 \in U_1 \cap \mathring{\mathbb{H}}$ and $x_2 \in U_2 \cap \partial\mathbb{H}$, such that $\alpha_1(x_1) = p = \alpha_2(x_2)$. But then if we let φ be the transition function, $\varphi = \alpha_2^{-1} \circ \alpha_1 : \alpha_1^{-1}(V_1 \cap V_2) \to \alpha_2^{-1}(V_1 \cap V_2)$. This means that $\varphi(x_1) = \alpha_2^{-1}(p) = x_2$.

But the transition function is a diffeomorphism and therefore its derivative $D\varphi(x_1)$ should be full rank (non-degenerate), by the inverse function theorem this means that there exists a neighbourhood $V \subset \mathring{\mathbb{H}}^d$ containing x_1 such that $\varphi(V)$ is open in $\mathbb{R}^d$ but $\varphi(V) \subset \mathbb{H}^d$ and $x_2 \in \partial\mathbb{H}^d \cap \varphi(V)$ such that any open ball $B_\varepsilon(x_2)$ contains points $y \in \mathbb{R}^d - \mathbb{H}^d \notin \varphi(V)$ and thus we reach a contradiction. □

This lemma is more powerful than it may look, we have now shown that a manifold with boundary is a union of two disjoint objects which we can now study independently.

Theorem 3.5 (Boundary of a Manifold is a Manifold): Let M be a d-manifold with boundary. Then ∂M is a $d-1$-manifold without boundary.

Proof: Let M be a d-manifold with boundary and $p \in \partial M \subset M$ with $\alpha : U \to V \subset M$ a coordinate patch around p. Then the restriction $\alpha|_{U \cap \partial\mathbb{H}^d} : U \cap \partial\mathbb{H}^d \to \alpha(U \cap \partial\mathbb{H}^d)$ is a diffeomorphism between the set $U \cap \partial\mathbb{H}^d$ open in $\mathbb{R}^{d-1}$ and the image set which can be written as $V \cap \partial M$ by Lemma 3.4. Letting U_0 be the open subset of $\mathbb{R}^{d-1}$ such that $U_0 \times \{0\} = U \cap \partial\mathbb{H}^d$ and letting V_0 be the image set of the above restriction of α, $V_0 = V \cap \partial M$ then the function $\alpha_0 : U_0 \to V_0$ defined by $\alpha_0(x) = \alpha(x, 0)$ is a diffeomorphism onto an open subset of ∂M from an open subset U_0 of $\mathbb{R}^{d-1}$. Finally since the derivative $D\alpha_0(x)$ is just the derivative $D\alpha(x, 0)$ ($D\alpha$ with the last column removed) is rank $d-1$ for all $x \in U_0$. Therefore α is a coordinate patch onto ∂M such that ∂M is locally diffeomorphic to open subsets of $\mathbb{R}^{d-1}$. Thus ∂M is a $d-1$-manifold without boundary.

Theorem 3.6 (Smooth Preimage of $\mathbb{R}^+$ is a Manifold with Boundary): Let U be an open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ a smooth function. Now let $M = \{x \in U \mid f(x) = 0\} = f^{-1}(\{0\})$ be the set of zeros of f. Suppose also that $Df(x)$ is rank-1 for all $x \in M$. Then $N = \{x \in U \mid f(x) \geq 0\} = f^{-1}([0, \infty))$ is a manifold with boundary and $\partial N = M$ is also a manifold.

Proof: Let p be a point in N, we will separate the cases where $f(p)$ is greater than 0 and equal to 0 so that we can deal with the boundary on its own.

1) f(p) > 0: Let $A = f^{-1}((0, \infty)) \subseteq U$, then the identity mapping $\alpha : A \to A$ is trivially a coordinate patch around $p \in N$ as A is open in $\mathbb{R}^n$.

2) f(p) = 0 Since $Df(p) \neq 0$, we have that at least one of the partial derivatives $\frac{\partial f}{\partial x_i}$ is non-zero. Without loss of generality we can assume that the partial derivative $i = n$ is non-zero and then we define a function $F : U \to \mathbb{R}^n$ as $F(x) = (x_1, x_2, \ldots, x_{n-1}, f(x))$ which is the identity for the first $n-1$ components and $f(x)$ for the nth component. Then its derivative is

$$DF(p) = \begin{pmatrix} 1 & 0 & \dots & 0 & \frac{\partial f}{\partial x_1}(p) \\ 0 & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & \dots & & 1 & \vdots \\ 0 & \dots & & 0 & \frac{\partial f}{\partial x_n}(p) \end{pmatrix}$$

which is full rank (also non-singular). Then by the inverse function theorem there exists an open neighbourhood $\tilde{U} \subseteq \mathbb{R}^n$ of p and an open set $\tilde{V} \subseteq \mathbb{R}^n$ such that $F|_{\tilde{U}} : \tilde{U} \to \tilde{V}$ is a diffeomorphism. Furthermore, F carries the open set $(\tilde{U} \cap N) \subseteq N$ onto the open set $(\tilde{V} \cap \mathbb{H}^n) \subseteq \mathbb{H}^n$ since $x \in N$ if and only if $f(x) \geq 0$. In particular this means that the restriction $F_{\tilde{U} \cap M}$ carries $(\tilde{U} \cap M) \subseteq M$ onto $(\tilde{V} \cap \partial\mathbb{H}^n)$ since $x \in M$ if and only if $f(x) = 0$. This means that the restriction $F^{-1}_{\tilde{V} \cap \mathbb{H}^n} : \tilde{V} \cap \mathbb{H}^n \to \tilde{U} \cap N$ is the required coordinate patch on N making N an n-manifold with boundary. And the more strict case $F^{-1}_{\tilde{V} \cap \partial\mathbb{H}^n} : \tilde{V} \cap \partial\mathbb{H}^n \to \tilde{U} \cap M$ is a coordinate patch on M since it maps points of the form $(p_1, p_2, \dots, p_{n-1}, 0) \in \tilde{V}$ into N. $\square$

Definition: For a d–manifold with or without boundary. A coordinate patch α around a point $p \in M$, together with its domain U open in $\mathbb{H}^d$ or $\mathbb{R}^d$ can be collectively referred to as a **chart around** p and is denoted $\{\alpha, U\}$.

The union of all charts covering M is referred to as an **atlas on** M and is denoted usually by some fancy letter $\mathcal{A}(M) = \bigcup_{i \in I} \{\alpha_i, U_i\}$.

4 Tangent Vectors and Smooth Mappings Between Manifolds

In this section of the course we begin to work with the relation of vector spaces to manifolds and smooth functions on manifolds. This will involving proving some very powerful theorems for constructing manifolds.

Definition (Tangent Vector to a point in $\mathbb{R}^n$): Given a point $x \in \mathbb{R}^n$, we define a tangent vector to $\mathbb{R}^n$ at x to be a pair $(x; \mathbf{v})$ where $\mathbf{v} \in \mathbb{R}^n$. The set of all tangent vectors to $\mathbb{R}^n$ at x forms a vector space under the usual operations of vector addition and we refer to it as the **tangent space to $\mathbb{R}^n$ at** x and denote it by $T_x\mathbb{R}^n$.

Definition (Tangent Vector to a Manifold): Let $M \subseteq \mathbb{R}^n$ be a d-manifold with boundary. Given a point $p \in M$ and a coordinate patch $\alpha : U \to V$ around p. There exists a point $x_0 \in U \subseteq \mathbb{R}^d$ such that $\alpha(x_0) = p$. Then the **tangent space of** M **at** p is the vector space spanned by the columns of $D\alpha(x_0)$ and we denote the space by T_pM.

Said differently, the vectors $(\frac{\partial \alpha_1}{\partial x_j}, \frac{\partial \alpha_2}{\partial x_j}, \dots, \frac{\partial \alpha_n}{\partial x_j})\big|_{x_0} \in \mathbb{R}^n$, where $j = 1, \dots, d$ are tangent vectors to M at p and they form a linearly independent basis for the space T_pM.

Lemma 4.1: For a d-Manifold with boundary, the definition of T_pM does not depend on the coordinate patch α.

In lectures the proof was left as an exercise so this is only an attempt which I'm pretty confident is correct.

Proof: We can actually prove this for a manifold with or without boundary. Let M be a d-manifold (with or without boundary) in $\mathbb{R}^n$ and arbitrarily choose some point $p \in M$ which can have any number of valid coordinate patches into a neighbourhood around it from U open in $\mathbb{R}^d$ or $\mathbb{H}^d$. From the definition we can say that for a coordinate patch α around p with $\alpha(x_0) = p$, the space T_pM is composed of the vectors (columns) in $D\alpha(x_0)$ which specifically take the form:

$$\left(\frac{\partial \alpha_1}{\partial x_j}, \frac{\partial \alpha_2}{\partial x_j}, \ldots, \frac{\partial \alpha_n}{\partial x_j}\right)\bigg|_{x_0} \in \mathbb{R}^n \quad \text{where} \quad j = 1, \ldots, d.$$

Since α is a coordinate patch, it is true that $D\alpha$ has rank$-d$ and therefore all d of these vectors are linearly independent and hence form a basis of the space T_pM. Furthermore since $\mathbb{R}^d$ is spanned by the orthonormal set of standard basis vectors $(e_1, e_2, \ldots, e_d)$, the space T_pM is spanned by the projections of this basis along the Euclidean axes:

$$(D\alpha(x_0) \cdot e_j) = \left(\frac{\partial \alpha}{\partial x_j}\right).$$

So for any α that is a coordinate patch around p, it will have a rank$-d$ derivative at p composed of d independent vectors which form a basis for T_pM. And this basis can always be projected along orthogonal directions of $\mathbb{R}^d$ thereby spanning the space T_pM. $\square$

Proposition 4.2: Let M, N be manifolds with boundary and $f : M \to N$ a smooth function. Then for some extension $\tilde{f} : \tilde{M} \to \tilde{N}$ which is also smooth, $Df(p) = D\tilde{f}(p)$ defines a linear map $T_pM \to T_{f(p)}N$ for all $p \in M$.

Proof: Let $\alpha_1 : U_1 \to V_1 \subseteq M$ and $\alpha_2 : U_2 \to V_2 \subseteq N$ be charts around p and $f(p)$ respectively. By shrinking U_1 and consequentially V_1, we can assume that $f(V_1) \subseteq V_2$ and hence $\alpha_2^{-1} \circ f \circ \alpha_1 : U_1 \to U_2$ is well defined and smooth as a composition of smooth functions.

Now let $\tilde{f}$ be any smooth extension of f, then $\alpha_2^{-1} \circ f \circ \alpha_1 = \alpha_2^{-1} \circ \tilde{f} \circ \alpha_1 = \varphi$ since the action of α_1 restricts the domain of the composition to U_1 on which $\tilde{f}$ agrees with f. Thus if we apply α_2 on the left of this equation we find that $\tilde{f} \circ \alpha_1 = \alpha_2 \circ \varphi$ does not depend on $\tilde{f}$ and taking derivatives of both sides at p by the chain rule, we obtain:

$$D\tilde{f}(p) \circ \underbrace{D\alpha_1(\alpha_1^{-1}(p))}_{Image=T_pM} = \underbrace{D\alpha_2(\alpha_2^{-1}(f(p))}_{Image=T_{f(p)}N} \circ D\varphi(\alpha_1^{-1}(p)).$$

Which shows that when $D\tilde{f}(p)$ is restricted to T_pM, it does not depend on $\tilde{f}$ and the image of this restriction is contained within $T_{f(p)}N$, namely $D\tilde{f}(p)|_{T_pM} \subseteq T_{f(p)}N$. $\square$

Definition (Tangent Bundle): For $M \subseteq \mathbb{R}^m$, the **tangent bundle of** M is defined as the union of all tangent spaces T_pM for $p \in M$ denoted by $TM = \bigcup\limits_{p\in M} T_pM$.

In set notation the tangent bundle of M can be written as:

$$TM = \{(x, \mathbf{v}) \in M \times \mathbb{R}^n \mid \mathbf{v} \in T_pM\} \subset \mathbb{R}^{n+m}$$

Proposition 4.3 (The tangent bundle is a Manifold): Let $M \subseteq \mathbb{R}^n$ be a d-manifold, then the tangent bundle TM of M is a $2d$-manifold with boundary.

Proof: Take some point $(p, \nu) \in M \times T_pM$ and let $\alpha : U \to M$ be a chart around p mapping the open subset $U \subset \mathbb{H}^d$ into M. Now define the function

$$\beta : (\mathbb{R}^d \times U) \subset \mathbb{H}^{2d} \to \mathbb{R}^{n+m} \equiv \beta(x, y) = (\alpha(y), D\alpha(y) \cdot x).$$

Then β is a smooth function in y because α is smooth and in x because its linear in y. We can then consider the composition of β with the function $\varphi(x, y) = (D\alpha^{-1}(x) \cdot y, \alpha^{-1}(x))$:

$$(\varphi \circ \beta)(x, y) = \varphi(\alpha(y), D\alpha(y) \cdot x) = (D\alpha^{-1}(\alpha(y)) \cdot D\alpha(y) \cdot x, \alpha^{-1}(\alpha(y))).$$

We can group together the first derivative by the chain rule which gives:

$$(D(\alpha \circ \alpha^{-1})(y) \cdot x, (\alpha^{-1} \circ \alpha)(y)) = (x, y),$$

thus $\beta \circ \varphi$ is the identity mapping $Id_{\mathbb{R}^{n+m}}$ which is a diffeomorphism and thus the two functions including in the composition are also diffeomorphisms which implies that β is a diffeomorphism. Now we must show that β has a rank$-2d$ derivative at all $(x, y) \in \mathbb{R}^d \times U$. We can compute the derivative in matrix form as follows:

$$D\beta(x, y) = \begin{pmatrix} 0 & \dots & 0 & \partial_1\alpha_1(y) & \dots & \partial_d\alpha_1(y) \\ 0 & \dots & 0 & \partial_1\alpha_2(y) & \dots & \partial_d\alpha_2(y) \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \partial_1\alpha_n(y) & \dots & \partial_d\alpha_n(y) \\ \partial_1\alpha_1(y) & \dots & \partial_d\alpha_1(y) & & & \\ \vdots & & \vdots & & (n \times d) & \\ \partial_1\alpha_n(y) & \dots & \partial_d\alpha_n(y) & & & \end{pmatrix}$$

This matrix is clearly rank-2d because $D\alpha(y)$ is rank-d and the upper left $n \times d$ block is the zero matrix. Therefore $\{\beta, \mathbb{R}^d \times U\}$ is a chart around $(x, y) \in TM$ and thus TM is a $2d$-manifold with boundary in $\mathbb{R}^{n+m}$.

Proposition 4.4 (Derivatives of Smooth Maps between Manifolds): Let $M, N, L \subset \mathbb{R}^n$ be d-manifolds with boundary.

1. Let $f : M \to N$ be a smooth map between M and N. Then the differential map $Df : TM \to TN$ is also smooth.

2. Let $f : M \to N$, $g : N \to L$ be smooth maps, then $D(g \circ f) = Dg \circ Df$

Proof (1): Given a d-manifold M, we know that $f : M \to \mathbb{R}^m$ is smooth iff for any chart $\{\alpha, U\}$ around $p \in M$, the composition $f \circ \alpha : U \to \mathbb{R}^m$ is smooth since charts are diffeomorphisms. Given some $(p, \nu) \in TM$ we can define the chart $\beta : \mathbb{R}^d \times U \to \mathbb{R}^{n+m}$ around (p, ν) as in the previous proof:

$$\beta(x, y) = (\alpha(y), D\alpha(y) \cdot x).$$

Let f be smooth then the derivative of the composition $f \circ \beta$ is also smooth and we have:

$$D(f \circ \beta)(x, y) = Df\big(\alpha(y), D\alpha(y) \cdot x\big)$$

$$= \big(f \circ \alpha(y), Df(\alpha(y))D\alpha(y) \cdot x\big) = \big((f \circ \alpha)(y), D(f \circ \alpha)(y) \cdot x\big).$$

Which is smooth (in y) because $f \circ \alpha$ is and (in x) because its linear in x. $\square$

Proof (2): Let $f : M \to N$, $g : N \to L$ be smooth maps between the three manifolds and $(p, \nu) \in TM$. Then we simply apply the chain rule:

$$\begin{aligned} D(g \circ f)(p, \nu) &= \big((g \circ f)(p), D(g \circ f)(p) \cdot \nu\big) \\ &= \big(g(f(p)), Dg(f(p)) \cdot Df(p) \cdot \nu\big) \\ &= Dg(f(p), Df(p) \cdot \nu) = Dg(Df(p, \nu)) = (Dg \circ Df)(p, \nu). \end{aligned}$$

$\square$

Definition (Regular Value of a Function): Let $f : M \to N$ be a smooth function between manifolds. Then we say $p \in N$ is a **regular value of** f if $Df : T_xM \to T_pN$ is surjective (onto) for all $x \in f^{-1}(p)$, otherwise we say that p is a **critical value**.

Theorem 4.5 (Preimage Theorem): Let $f : M \to N$ be a smooth function between two manifolds. If $p \in N$ is a regular value of f then the preimage $L = f^{-1}(p)$ is a manifold. Moreover, $T_xL = \ker(Df(x) : T_xM \to T_pN)$.

Proof: Let $f : M \to N$ be a smooth map between (d_1 and d_2)-manifolds and let $L = f^{-1}(p)$ where $p \in N$. Let $x_0 \in L$ and $\alpha_1 : U_1 \to V_1 \subset M$ and $\alpha_2 : U_2 \to V_2 \subset N$ be charts around x_0 and p respectively. We can always shrink U_1 and consequentially V_1 such that $f(V_1) \subset V_2$ such that $\varphi \equiv \alpha_2^{-1} \circ f \circ \alpha_1 : U_1 \to U_2$ is a smooth function. Suppose p is a regular value of f, then $D\varphi(x_0)$ is surjective and we can find an invertible matrix B such that

$$\begin{pmatrix} D\varphi(x_0) \\ B \end{pmatrix}$$

is invertible and full rank. Then we can construct another function $\psi : U_1 \to U_2 \times \mathbb{R}^{d_1-d_2}$, $x \mapsto (\varphi(x), Bx)$ whose derivative is this invertible matrix. Then by the inverse function theorem there exists a subset $p \ni \tilde{U}_1 \subset U_1$ and $U_2 \supset \tilde{U}_2 = \psi(\tilde{U}_1)$ such that $\psi|_{\tilde{U}_1}$ is a diffeomorphism. This implies that $\psi^{-1}|_{\tilde{U}_2}$ is smooth. Now set $\alpha_L = (\alpha_1 \circ \psi^{-1})|_{\tilde{U}_2 \cap \{\alpha_2^{-1}(p)\} \times \mathbb{R}^{d_1-d_2}}$, this only maps points (x, y) where x are points in $\tilde{U}_2$ whose image under α_2 is p along with points y in $\mathbb{R}^{d_1-d_2}$ onto points $q \in L \cap \alpha_1(\tilde{U}_1) \subset \mathbb{R}^{d_1}$. Clearly α_L is a chart for L around x_0 and its derivative is $\mathrm{rank}(d_1-d_2)$ so we conclude that L is a (d_1-d_2)−manifold.

Finally, since $f \circ \alpha_L(x) = p$ (constant image) for all $x \in \tilde{U}_2 \cap \{\alpha_2^{-1}(p)\} \times \mathbb{R}^{d_1-d_2}$ and $D(f \circ \alpha_L) = Df \circ D\alpha_L = 0$ we have that $im(D\alpha_L) \subseteq \ker Df$ and since both of these vector spaces have the same dimension we find that the two spaces are actually equal and therefore $T_x L = \ker(Df(x) : T_x M \to T_p N)$ □

Theorem 4.6 (Sard): Let $f : M \to N$ be a smooth map between manifolds. Then the set of values $crit(f) \subseteq N$ has measure zero (is a null set). i.e. For $N \subset \mathbb{R}^n$ the set $\{x \in M \mid rank(Df(x)) < n\}$ has measure zero.

Theorem 4.7 (Preimage Theorem for Manifolds with Boundary): Let $f : M \to N$ be a smooth function between manifolds with boundary. Suppose that $p \in N$ is a regular value of f and also a regular value of $f|_{\partial M} : \partial M \to N$, then $f^{-1}(p) = L$ is a manifold with boundary and $\partial L = f^{-1}(0) \cap \partial M$.

Proof: Let $f : M \to N$ be a smooth function between two manifolds with boundary $M \subset \mathbb{R}^m$ (of dimension d_1) and $N \subset \mathbb{R}^n$ (of dimension d_2). Let $x_0 \in L = f^{-1}(p)$. Now suppose p is a regular value of f and $f|_{\partial M}$. If $x_0 \in L$ is not in the boundary ∂M then as in the without-boundary version of the theorem we have that L is a submanifold (to M) in a neighbourhood of x_0. Therefore let $x_0 \in \partial M \cap f^{-1}(p)$. We can then pick charts $\alpha_1 : U_1 \to V_1 \subset M$ and $\alpha_2 : U_2 \to V_2 \subset N$ around x_0 and p respectively such that $\alpha_1^{-1}(x_0) = 0$ and $\alpha_2^{-1}(p) = 0$. Since these are manifolds with boundary we have that the domains of the charts are the open subsets $U_1 \subset \mathbb{H}^{d_1}$ and $U_2 \subset \mathbb{H}^{d_2}$ respectively. We can always shrink U_1 such that $f(V_1) \subset V_2$ such that the function $\varphi \equiv \alpha_2^{-1} \circ f \circ \alpha_1 : U_1 \to U_2$ is smooth.

Let W be an open subset of $\mathbb{R}^{d_1}$ such that $W \cap \mathbb{H}^{d_1} = U_1$ and let $\tilde{\varphi} : W \to \mathbb{H}^{d_2}$ be the smooth extension of φ over W. Now we may similarly shrink W small enough that its intersection with the half-plane is still U_1 but that $\tilde{\varphi}$ is regular on W. In other words we have shrunk the set to avoid potential critical points. Hence $A \equiv \tilde{\varphi}^{-1}(0)$ is a submanifold of $\mathbb{R}^{d_1}$ by the without boundary version of the preimage theorem. Now consider the real-valued function $F : \tilde{\varphi}^{-1}(0) \to \mathbb{R}$ mapping $(x_1, x_2, \ldots, x_{d_1}) \mapsto x_{d_1}$. $0 \in \mathbb{R}$ must be a regular value of F, since if it weren't, the tangent space $T_0 A$ would lie solely in points with $x_{d_1} = 0$ which contradicts p being a regular value of $f_{\partial M}$. Hence $F^{-1}(0) = f^{-1}(0) \cap \partial M$ is itself a manifold without boundary and is the boundary of L denoted ∂L. □

Although some of these proofs may be tedious and its true that they are, they can now be used to construct a manifold in a much easier way than proving all four conditions of a coordinate patch. The results will also be implicit and helpful in our important subsequent discussions of the most geometric section of the course.

5 Multilinear Algebra and Tensors

Discussion (Motivation for Stokes' Theorem): Upon inspection of the following laws:

$$\int_a^b f'(x)\ dx = f(b) - f(a) \qquad \text{Fundamental Theorem of Calculus}$$

$$\int\int_V\int \mathrm{div}\vec{E} = \int\int_{\partial V} \vec{E}\cdot d\vec{n} \qquad \text{Gauss' Law}$$

$$\int\int_S \left(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}\right)\ dxdy = \int_{\partial S}(A\ dx + B\ dy) \qquad \text{Stokes' Law on the Plane}$$

We can see that mathematically they all have the same basic principle of exchanging an integral over a whole object for an integral only over its boundary. It turns out that indeed, all three of these equations are special cases of the general Stokes' theorem which takes the form

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

But what kind of object is ω that it can be integrated? Well the fact it can be integrated means that we can measure the volumes of infinitesimally small cubes (or parallelopipeds) locally that make up ω. What we mean is that if we let (C_i) be a collection of cubes making up a manifold M, we can imagine integrating ω over M as a function of these cubes:

$$\int_M \omega = \lim_{\substack{C_i \text{ cubes} \\ \text{size}(C_i)\to 0}} \sum_i \omega(C_i).$$

Since we only care about these objects locally, we are only interested in restrictions of ω such as ω_p : pipeds in $T_pM \to \mathbb{R}$. The reason for this is that a parallelopiped in T_pM is spanned by vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d \in T_pM$ so we can write ω_p in terms of these vectors for each point $p \in M$. Since volume scales with extensive variables we also must have that ω_p is a homogeneous degree-1 function of these vectors i.e. $\omega(\lambda\mathbf{v}_1, \lambda\mathbf{v}_2, \ldots, \lambda\mathbf{v}_d) = \lambda\omega(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d)$ and in particular $\omega(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d) = -\omega(-\mathbf{v}_1, -\mathbf{v}_2, \ldots, -\mathbf{v}_d)$. Functions satisfying the properties of this object ω are known as **differential forms**.

With this discussion we now develop some concepts in multilinear algebra so we can get back to the exciting prospect of differential forms and the putting so many important equations under one mathematical basis with solid foundations.

Definition (Multilinearity): Let V be a vector space and let V^k denote the set of all k-tuples of vectors in V. A function $T : V^k \to \mathbb{R}$ is said to be **linear in the i-th variable** if, given a fixed set of vectors $\mathbf{v}_j$, where $j \neq i$, the mapping $\mathbf{v} \mapsto T(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_d)$ is linear. Then T is said to be **multilinear or a k-tensor on V** if it is linear in the i-th variable for all $i = 1, \ldots, k$.

We denote the set of all k-tensors on V by the symbol $\mathcal{L}^k(V)$. If $k = 1$ then $\mathcal{L}^1(V)$ is just the set of all linear transformations $T : V \to \mathbb{R}$ and we usually refer to this as the **dual space of** V and denote it by V^*.

Proposition 5.1 ($\mathcal{L}^k(V)$ is a Vector Space): Let V be a vector space. Then $\mathcal{L}^k(V)$ is a vector space if we define the operations of addition and scalar multiplication as $f, g \in \mathcal{L}^k(V)$ and $\lambda \in \mathbb{R}$:

$$(f+g)(\mathbf{v}_1, \ldots, \mathbf{v}_k) = f(\mathbf{v}_1, \ldots, \mathbf{v}_k) + g(\mathbf{v}_1, \ldots, \mathbf{v}_k),$$

$$(\lambda f)(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \lambda(f(\mathbf{v}_1, \ldots, \mathbf{v}_k))$$

Proof: We simply check that the axioms defining a vector field are satisfied. For $f, g, h \in \mathcal{L}^k(V)$ and $c, d \in \mathbb{R}$ for simplicity of writing we will use $v \in V^k$ to denote any k-tuple of vectors in V.

- Commutativity of addition: $(f+g)(v) = f(v) + g(v) = g(v) + f(v) = (g+f)(v)$ which holds since these functions are real-valued.
- Transitivity of addition: $(f + (g+h))(v) = f(v) + (g+h)(v) = f(v) + g(v) + h(v) = (f+g)(v) + h(v) = ((f+g)+h)(v)$ which follows from commutativity.
- Zero Element: By definition, the zero element is the k-tensor whose value is zero for every k-tuple of vectors in V if we denote it by $o : V^k \to \mathbb{R}$ we have that $(f+o)(v) = f(v) + 0 = f(v)$.
- Additive inverse: $f(v) + (-1)f(v) = 0 = (f + (-f))(v)$.
- Multiplicative Identity: The real number 1 acts as a multiplicative identity since $(1 \cdot f)(v) = 1(f(v)) = f(v)$.
- Compatibility of multiplication: $(c(d \cdot f)(v)) = cd(f(v)) = (cd \cdot f)(v)$.
- Distributivity (1): $((c+d) \cdot f)(v) = (c+d)f(v) = cf(v) + df(v)$ again holds since f is real valued.
- Distributivity (2): $(c(f+g))(v) = c((f+g)(v)) = c(f(v) + g(v)) = cf(v) + cg(v)$.

Since all axioms are satisfied we conclude that $\mathcal{L}^k(V)$ is a vector space. $\square$

Lemma 5.2 (Uniqueness of k-tensors): Let $a_1, a_2, \ldots, a_n$ be a basis for the vector space V. If $f, g : V^k \to \mathbb{R}$ are k-tensors on V, and if

$$f(a_{i_1}, \ldots, a_{i_k}) = g(a_{i_1}, \ldots, a_{i_k})$$

for every k-tuple $I = (i_1, \ldots, i_k)$ of integers from the set $\{1, \ldots, n\}$, then $f = g$.

Proof: Given a k-tuple $(v_1, \ldots, v_k)$ of vectors in V, we use that fact that we can write a vector as a sum over the basis elements to write

$$v_i = \sum_{j=1}^{n} c_{ij} a_j \ .$$

Then using this to rewrite the image of the k-tuple under the k-tensor f

$$f(v_1, \ldots, v_k) = \sum_{j_1=1}^{n} c_{1j_1} f(a_{j_1}, v_2, \ldots, v_k)$$

$$\sum_{j_1=1}^{n} \sum_{j_2=1}^{n} c_{1j_1} c_{2j_2} f(a_{j_1}, a_{j_2}, v_3, \ldots, v_k).$$

We can apply these steps repeatedly until we have written the function as a linear combination of f on the basis elements of V^k

$$f(v_1, \ldots, v_k) = \sum_{1 \leq j_1, \ldots j_k, \leq n} c_{1j_1} c_{2j_2} \cdots c_{kj_k} \ f(a_{j_1}, \ldots a_{j_k}).$$

The same computation can be carried out on g and it follows that f and g agree on all k-tuples of vectors in V if they agree on all k-tuples of the basis elements. $\square$

Recall that in linear algebra we could completely determine the action of a linear transformation once its effects on the basis of a vector space were determined. Well we can observe that an identical principle holds for multilinear transformations once its basis elements are appropriately defined.

Definition (Elementary k-tensors): Let V be a vector space with basis $a_1, \ldots, a_n$. Let $I = (i_1, \ldots, i_k)$ be a k-tuple of integers from the set $\{1, \ldots, n\}$. There is a **unique** (by the preceding Lemma) k-tensor e^I on V such that, for every k-tuple $J = (j_1, \ldots, j_k)$ from $\{1, \ldots, n\}$ the action of e^I on the basis elements is given by

$$e^I(a_{j_1}, \ldots, a_{j_k}) = \begin{cases} 0 & \text{if } I \neq J \\ 1 & \text{if } I = J \end{cases} .$$

In particular we define the action of e^I by

$$e^I(v_1, \ldots, v_k) = e^{i_1}(v_1) \cdot e^{i_2}(v_2) \cdots e^{i_k}(v_k). \tag{5.1}$$

The tensors e^I are called the **elementary k-tensors** on V corresponding to the basis $a_1, \ldots, a_n$. When $k = 1$, the basis for V^* formed by the elementary tensors $e^1, \ldots, e^n$ is called the **dual basis** for V^*.

We can now prove that these tensors do in fact exist and form a basis for $\mathcal{L}^k(V)$.

Theorem 5.3 (**e^I form a basis for $\mathcal{L}^k(V)$**): Let V be a vector space with basis $a_1, \ldots, a_n$. Let $\{I\} = \{i_1, \ldots, i_k\}$ be the set of all k-tuples of the integers $\{1, \ldots, n\}$, then the elementary k-tensors $\{e^I\}$ as defined above, exist and form a basis for $\mathcal{L}^k(V)$. In particular $\dim(\mathcal{L}^k(V)) = n^k$.

Proof: We first consider the case $k = 1$. We know that we can determine a linear transformation by specifying its values on the basis elements so we can define $e^i \in V^*$ by the equation

$$e^i(a_j) = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} .$$

Then these are the desired 1-tensors which form the basis of V^* **dual** to V. In the case $k > 1$, we define e^I as in the equation above by

$$e^I(v_1, \ldots, v_k) = e^{i_1}(v_1) \cdot e^{i_2}(v_2) \cdots e^{i_k}(v_k).$$

Now since this definition of e^I is written as a product of elementary 1-tensors, it follows that each e^i is linear and since multiplication is distributive we have that e^I is multilinear of order k. Furthermore, any term $e^i(a_j)$ in the product vanishes if $i \neq j$ so one can immediately see that e^I takes the appropriate value on each k-tuple of basis vectors:

$$e^I(a_{j_1}, \ldots, a_{j_k}) = \begin{cases} 0 & \text{if } I \neq J \\ 1 & \text{if } I = J \end{cases}$$

Now to show that e^I forms a basis for $\mathcal{L}^k(V)$, we show that given a k-tensor f on V, we can write f uniquely as a linear combination of the tensors e^I. For each k-tuple $I = \{1, \ldots, n\} \in \{I\}$, let s_I be the scalar defined by

$$s_I = f(a_{i_1}, \ldots, a_{i_k}).$$

Then consider the tensor formed by the summation

$$g = \sum_J s_J e^J$$

where J is any k-tuple of integers from the set $\{1, \ldots, n\}$. Then the value of g on the k-tuple $(a_{i_1}, \ldots, a_{i_k})$ is equal to s^I by the definition of e^I since I must be equal to J for a non-vanishing contribution. Then by the previous lemma $f = g$ and hence the set $\{e^I\}$ span the set $\mathcal{L}^k(V)$ and form a unique representation of all k-tensors.

Lastly since there are n^k distinct k-tuples from the set $\{1,\dots,n\}$, there must be n^k linearly independent elementary k-tensors spanning $\mathcal{L}^k(V)$. So $\dim(\mathcal{L}^k(V)) = n^k$. $\square$

Definition (The Tensor Product): Let f be a k-tensor on V and let g be an l-tensor on V. We define a $k+l$ tensor $f \otimes g$ on V by the equation

$$(f \otimes g)(v_1,\dots,v_{k+l}) = f(v_1,\dots,v_k)\cdot g(v_{k+1},\dots,v_{k+l}).$$

The function $f \otimes g$ is multilinear, and its called the **tensor product** of f and g.

Theorem 5.4 (Properties of Tensor Product): Let f,g,h be tensors on V. Then the following properties hold:

1. **(Associativity):** $f \otimes (g \otimes h) = (f \otimes g) \otimes h$.
2. **(Homogeneity):** $(cf) \otimes g = c(f \otimes g) = f \otimes (cg)$
3. **(Distributivity):** If f and g have the same order then

$$(f+g) \otimes h = f \otimes h + g \otimes h$$

$$h \otimes (f \otimes g) = h \otimes f + h \otimes g$$

4. Given a basis $a_1,\dots,a_n$ for V, the corresponding elementary tensors e^I satisfy the equation

$$e^I = e^{i_1} \otimes e^{i_2} \otimes \cdots \otimes e^{i_k}$$

where $I = (i_1,\dots,i_k)$. The proofs of these properties are very straightforward using the definition of the tensor product so I will omit them here.

Definition (Dual Transformation): Let $A : V \to W$ be a linear transformation between vector spaces. We define the dual transformation $A^* : \mathcal{L}^k(W) \to \mathcal{L}^k(V)$ as follows. If $f \in \mathcal{L}^k(W)$, and if $v_1,\dots,v_k$ are vectors in V, then

$$(A^* f)(v_1,\dots,v_k) = f(A(v_1),\dots,A(v_k)).$$

The dual transformation $A^* f$ can be viewed as the composition of $A \times \cdots \times A$ and f on any k-tuple of vectors in V. Note that A^* itself goes in the reverse direction of A and since f is multilinear and A is linear we can show that A^* is linear as a map between tensors.

Theorem 5.5 (Properties of the Dual Transformation): Let $A : V \to W$ be a linear transformation and $A^* : \mathcal{L}^k(W) \to \mathcal{L}^k(V)$ be the dual transformation. Then the following properties hold:

1. A^* is linear.

2. $A^*(f \otimes g) = A^* f \otimes A^* g$
3. If $B : W \to X$ is a linear transformation then $(B \circ A)^* f = A^*(B^* f)$.

Proof: These proofs are once again straightforward but this time I won't spare the detail as the dual transformation is a bit more abstract than the tensor product.

(1): Let f, g be k-tensors on W, then we can prove linearity by computing

$$(A^*(af + bg))(v_1, \ldots, v_k) = (af + bg)(A(v_1), \ldots, A(v_k)).$$

Then since multiplication is distributive this reduces to

$$= af((A(v_1), \ldots, A(v_k))) + bg(A(v_1), \ldots, A(v_k)) = (aA * f + bA^* g)(v_1, \ldots, v_k).$$

Hence T^* is linear.

(2): Similarly we prove distributivity over the tensor product $f \otimes g$ by considering f as a k-tensor and g as an l-tensor and computing

$$A^*(f \otimes g)(v_1, \ldots, v_{k+l}) = (f \otimes g)(A(v_1), \ldots, A(v_{k+l})).$$

then by the definition of the tensor product, this becomes

$$f(A(v_1), \ldots, A(v_k)) \cdot g(A(v_{k+1}), \ldots, A(v_{k+l}))$$

$$= [A^* f(v_1, \ldots, v_k)] \cdot [A^* g(v_{k+1}, \ldots, v_{k+l})] = (A^* f \otimes A^* g)(v_1, \ldots, v_{k+l}).$$

(3): This property is also shown by computation, let $A^* : \mathcal{L}^k(W) \to \mathcal{L}^k(V)$ and $B^* : \mathcal{L}^k(X) \to \mathcal{L}^k(W)$ be dual transformations. Then

$$(B \circ A)^* f(v_1, \ldots, v_k) = f((B \circ A)(v_1), \ldots, (B \circ A)(v_k)) = f(B(A(v_1)), \ldots, B(A(v_k))).$$

It follows that

$$= B^* f(A(v_1), \ldots, A(v_k)) = A^*(B^* f)(v_1, \ldots, v_k)$$

which concludes the proof. □

Definition (Permutations and the set S_k): Let $k \geq 2$. A **permutation** of the set of integers $\{1, \ldots, k\}$ is a one-to-one function σ mapping this set onto itself. We denote the set of all such permutations by S_k. The set S_k is closed under σ^{-1} and compositions and has $k!$ elements.

Definition (Elementary Permutation): Given $1 \leq i < k$, let $e_i \in S_k$ be the permutation defined by setting $e_i(j) = j$ for all j not equal to i or $i + 1$ and setting

$$e_i(i) = i + 1 \qquad \text{and} \qquad e_i(i + 1) = i.$$

Then we call e_i an elementary permutation and e_i is its own inverse. Basically an elementary permutation swaps two neighbouring integers and fixes all of the others.

Lemma 5.6 : Every $\sigma \in S_k$ is a composite of elementary permutations.

Proof: If σ is the identity permutation which fixes all elements of $\{1, \dots, k\}$ then the result holds since $e_j \circ e_j$ is the identity for any j.
We can show the proof for non-trivial permutations by induction. Suppose σ fixes the first $i-1$ integers in the set. Then since σ is one-to-one, $\sigma(i)$ must be a number different from $1, \dots, i-1$. In the case that $\sigma(i) = i$, letting $\sigma' = \sigma$ and π be the identity permutation proves the result for integers up to i and we can move onto the next induction step. In the case that $\sigma(i) = l > i$, we set

$$\sigma' = e_i \circ \cdots \circ e_{l-1} \circ \sigma$$

then σ' fixes the integers $\{1, \dots, i-1\}$ and actually fixes i aswell since $\sigma(i) = l$ and by the definition of elementary permutations we can write

$$\sigma'(i) = e_i(\cdots(e_{l-1}(\sigma(i)))\cdots) = e_i(\cdots(e_{l-1}(l))\cdots) = \cdots = e_i(i+1) = i$$

then our result holds by induction since we can continue to construct such composites for $i+1, i+2, \dots$ up to k. □

Definition (Sign of a Permutation): We define the **sign of** σ to be -1 if the number of inversions in σ is **odd** and $+1$ if the number is **even**. In either case we call σ an odd/even permutation. We denote the sign of a permutation by the function $\mathrm{sgn}(\sigma)$ which returns ± 1.

Lemma 5.7 (Properties of sgn(σ): Let $\sigma, \tau \in S_k$.

1. If it takes a composition of m elementary permutations to express σ, then $\mathrm{sgn}(\sigma) = (-1)^m$.
2. $\mathrm{sgn}(\sigma \circ \tau) = \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}(\tau)$.
3. $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$.
4. If $p \neq q$ and σ exchanges p and q but fixes all other integers, then $\mathrm{sgn}(\sigma) = -1$.

Proof: First we show a useful result which directly relates to the Lemma itself and that is that for any σ, we can show that

$$\mathrm{sgn}(\sigma \circ e_l) = -\mathrm{sgn}(\sigma)$$

where e_l is an elementary permutation. Given a permutation σ, let $\tau = \sigma \circ e_l$, then for the set of integers $\{1, \dots, l-1, l, \dots k\}$ with which we are concerned, we have that

$$(\sigma(1), \dots, \sigma(l+1), \sigma(l), \dots, \sigma(k)) = (\tau(1), \dots, \tau(l), \tau(l+1), \dots, \tau(k)).$$

If the integers $p \neq q$ are both $\neq l$ or $l+1$, then $\sigma(p) = \tau(p)$ and $\sigma(q) = \tau(q)$ thus if p and q form an inversion in one sequence then they form the same inversion in the other. This is also true if only one of p or q are equal to either l or $l+1$ since one will be the same under both permutations and they will not swap order under e_l. However, if $\sigma(l)$ and $\sigma(l+1)$ form an inversion in one sequence then they do not under τ and vice versa. Hence the sequence formed by τ has either one more or one less inversion than the one formed by σ.

Now we prove the main results.

1. The identity permutation has sign +1, and by the previous step, the composition of each m required elementary permutations will change the sign m times.

2. Let σ be composed of m elementary permutations and τ be composed of n elementary permutations. Then $\sigma \circ \tau$ is composed of $(m+n)$ elementary permutations and the result follows from the equation $(-1)^{m+n} = (-1)^m(-1)^n$.

3. Since $\sigma^{-1} \circ \sigma$ is the identity permutation, we must have that $(\mathrm{sgn}(\sigma^{-1}))(\mathrm{sgn}(\sigma)) = +1$.

4. For elements $q < p$ of an increasing sequence, we can always write that $q = p + l$ where $l = 0, 1, \ldots$ then the number of inversions generated by swapping p and q will always be $2l - 1$ by basic counting. Thus a permutation which completes this task is odd. □

Definition (Permutations on Tensors): Let f be a k-tensor in $\mathcal{L}^k(V)$, and let σ be a permutation of $\{1, \ldots, k\}$, then we define f^σ by the equation

$$f^\sigma(v_1, \ldots, v_k) = f(v_{\sigma(1)}, \ldots, v_{\sigma(k)}).$$

For any elementary permutation e, we say that f is **symmetric** if $f^e = f$ and **alternating** if $f^e = -f$. We shall use $S^k(V)$ to denote the set of all symmetric k-tensors on V and $\mathcal{A}^k(V)$ to denote the set of all alternating tensors on V. Note that $\mathcal{A}^1(V) = \mathcal{L}^1(V) = V^*$.

Lemma 5.8 (Linearity of Permutations): Let $f \in \mathcal{L}^k(V)$, and let $\sigma, \tau \in S_k$. The transformation $f \to f^\sigma$ is a linear transformation $\mathcal{L}^k(V) \to \mathcal{L}^k(V)$. It has the property that

$$(f^\sigma)^\tau = f^{\tau \circ \sigma}$$

Proof: To prove linearity. Let $f, g \in \mathcal{L}^k(V)$ we compute

$$(af + bg)^\sigma(v_1, \ldots, v_k) = (af + bg)((v_{\sigma(1)}, \ldots, v_{\sigma(k)}).$$

Then since f and g are linear we write

$$= af(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) + bg(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = (af^\sigma + bg^\sigma)(v_1, \ldots, v_k).$$

As for the property of compositions of permutations we compute

$$((f^\sigma)^\tau)(v_1, \ldots, v_k) = f^\sigma((v_{\tau(1)}, \ldots, v_{\tau(k)}) = f(v_{\tau(\sigma(1))}, \ldots, v_{\tau(\sigma(k))}).$$

It follows that

$$= f(v_{(\tau\circ\sigma)(1)}, \dots, v_{(\tau\circ\sigma)(k)}) = f^{\tau\circ\sigma}(v_1, \dots, v_k).$$

Which shows our result. $\square$

Now we can make a more general statement about when a tensor is symmetric or alternating for arbitrary permutations based on the fact that we can write any permutation as a composite of elementary ones.

Proposition 5.9 (Condition for Alternating Tensor): Let f be a k-tensor on V and $\sigma \in S_k$. The tensor f is alternating iff $f^\sigma = (\text{sgn } \sigma)f$ for all σ. Additionally, f is alternating iff for $v_i = v_j$ with $i \neq j$ then $f(v_1, \dots, v_k) = 0$.

Proof: Given an arbitrary permutation σ we can decompose it as

$$\sigma = \sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_m$$

where each σ_i is an elementary permutation. Then

$$f^\sigma = f^{\sigma_1\circ\sigma_2\circ\cdots\circ\sigma_m} = ((\cdots(f^{\sigma_m})\cdots)^{\sigma_2})^{\sigma_1}.$$

By f being an alternating tensor we can then express this as

$$= (-1)^m f = (\text{sgn } \sigma)f.$$

Now suppose $v_i = v_j$ but $i \neq j$. Let τ be the permutation that exchanges i and j but fixes all other numbers. Then

$$f^\tau(v_1, \dots, v_k) = f(v_1, \dots, v_k),$$

since $v_i = v_j$, but f is also alternating so we must also have that

$$f^\tau(v_1, \dots, v_k) = -f(v_1, \dots, v_k) \tag{5.2}$$

since $(\text{sgn } \tau) = -1$ (Lemma 5.7). It follows that $f(v_1, \dots, v_k) = 0$. $\square$

Now we can move to find a basis of elementary alternating tensors in the same way we derived the elementary tensors e^I.

Lemma 5.10 (Uniqueness of Alternating Tensors): Let $a_1, \dots, a_n$ be a basis for V. If $f, g \in \mathcal{A}^k(V)$ satisfy the equation

$$f(a_{i_1}, \dots, a_{i_k}) = g(a_{i_1}, \dots, a_{i_k})$$

for every ascending k-tuple of integers $I = (i_1 < \cdots < i_k)$ from the set $\{1, \dots, n\}$, then $f = g$.

Proof: It suffices to prove that f and g have the same value on an arbitrary k-tuple $(a_{j_1}, \dots, a_{j_k})$ of basis elements as shown in Lemma 5.2 for $\mathcal{L}^k(V)$ of which $\mathcal{A}^k(V)$ is a subspace. Let $J = (j_1, \dots, j_k)$ be a k-tuple of integers from $\{1, \dots, n\}$.

If any $j_p = j_q$ then f and g are zero on this k-tuple since they are alternating. If instead all the indices are unique, then let σ be the permutation of $\{1, \dots, k\}$ such that $I = (j_{\sigma(1)}, \dots, j_{\sigma(k)})$ is ascending in order. Then

$$f(a_{i_1}, \dots, a_{i_k}) = f^\sigma(a_{j_1}, \dots, a_{j_k}) = (\text{sgn } \sigma) f(a_{j_1}, \dots, a_{j_k}).$$

The same computation holds for g and since f and g agree on $(a_{i_1}, \dots, a_{i_k})$, this implies that they agree also on the k-tuple $(a_{j_1}, \dots, a_{j_k})$. □

Theorem 5.11 (A basis of $\mathcal{A}^k(V)$): Let V be a vector space with basis $a_1, \dots, a_n$. Let $I = (i_1 < \dots < i_k)$ be an ascending k-tuple of integers from the set $\{1, \dots, n\}$. There is a unique alternating tensor Ψ^I on V such that for every ascending k-tuple $J = (j_1, \dots, j_k)$ from the set $\{1, \dots, n\}$,

$$\Psi^I(a_{j_1}, \dots, a_{j_k}) = \begin{cases} 0 & \text{if } I \neq J, \\ 1 & \text{if } I = J. \end{cases}$$

The tensors Ψ^I are called **elementary alternating tensors**, they form a basis for $\mathcal{A}^k(V)$. They can be written as a sum over permutations of elementary tensors e^I as

$$\Psi^I = \sum_{\sigma \in S_k} (\text{sgn } \sigma)(e^I)^\sigma$$

Proof: Uniqueness of such a tensor follows from Lemma 5.10. To prove that Ψ^I exist, let Ψ^I be defined by the formula given in the above theorem. Now we first show that Ψ^I is alternating. Let $\tau \in S_k$, then

$$(\Psi^I)^\tau = \sum_{\sigma \in S_k} (\text{sgn } \sigma)((e^I)^\sigma)^\tau = \sum_{\sigma \in S_k} (\text{sgn } \sigma)(e^I)^{\tau \circ \sigma}$$

$$= (\text{sgn } \tau) \sum_{\sigma \in S_k} (\text{sgn } (\tau \circ \sigma)(e^I)^{\tau \circ \sigma} = (\text{sgn } \tau)\Psi_I.$$

where we have used that $\tau \circ \sigma$ ranges over all of S_k as σ itself does.

To show that Ψ^I has the desired value on the basis of V, let $J = (j_1 < \dots < j_k)$ be an arbitrary ascending k-tuple from $\{1, \dots, n\}$. Then we have

$$\Psi^I(a_{j_1}, \dots, a_{j_k}) = \sum_{\sigma \in S_k} (\text{sgn } \sigma)\; e^I(a_{j_{\sigma(1)}}, \dots, a_{j_{\sigma(k)}}).$$

By the definition of Ψ^I the only term in this sum which survives is the one for which σ is the permutation such that $I = (j_{\sigma(1)}, \dots, j_{\sigma(k)})$. But since I and J are both ascending, this only occurs if $I = J$ and σ is the identity permutation in which case the value of the term is 1. If $I \neq J$ then all terms vanish. Therefore Ψ^I is the desired alternating

k-tensor on V which satisfies the definition given in the theorem.

Lastly we must show that Ψ^I form a basis for $\mathcal{A}^k(V)$. That is, given $f \in \mathcal{A}^k(V)$ we show that f can be written uniquely as a linear combination of Ψ^I tensors. Let d_I be the real number obtained by the equation

$$d_I = f(a_{i_1}, \ldots, a_{i_k}),$$

for each ascending k-tuple of integers $I = (i_1 < \cdots < i_k)$ from the set $\{1, \ldots, n\}$. Then consider the tensor g given by

$$g = \sum_{J \text{ ascending}} d_J \Psi^J.$$

Then the value of g on the k-tuple $(a_{i_1}, \ldots, a_{i_k})$ is simply d_I which is equal to the value of f on this k-tuple. Therefore any $f \in \mathcal{A}^k(V)$ can be written as a unique linear combination of Ψ^I tensors over all ascending k-tuples of the given basis for V. $\square$

Remark (The dimension of $\mathcal{A}^k(V)$): A good question to ask is how many unique elementary alternating tensors are there? Or rather, what is the dimension of the space $\mathcal{A}^k(V)$? Well given any k and any subset of k integers from $\{1, \ldots, n\}$, **there is exactly one ascending k-tuple of these elements**, and hence one corresponding elementary alternating tensor. Thus the number of basis elements for $\mathcal{A}^k(V)$ is the number of distinct sets of k integers that can be chosen from n integers. The binomial coefficient to be precise:

$$\dim(\mathcal{A}^k(V)) = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Theorem 5.12 (Dual Transformation of an Alternating Tensor): Let $A : V \to W$ be a linear transformation. If f is an alternating tensor on W, then $A^* f$ is an alternating tensor on V.

Proof: Let $f \in \mathcal{L}^k(W)$ and $A^* : \mathcal{L}^k(W) \to \mathcal{L}^k(V)$ then $A^* f \in \mathcal{L}^k(V)$ by definition. Now consider a k-tuple $v_{i_1}, \ldots, v_{i_k}$ of vectors in V where $I = (i_1, \ldots, i_k)$ is a k-tuple of integers from the set $\{1, \ldots, n\}$. The value of $A^* f$ on this k-tuple is

$$A^* f(v_{i_1}, \ldots, v_{i_k}) = f(A(v_{i_1}), \ldots, A(v_{i_k})).$$

Suppose any two vectors $v_{i_p} = v_{i_q}$ where $p \neq q$, then since A is a linear transformation this implies that $A(v_{i_p}) = A(v_{i_q}) \in W$ but f is alternating on W so in this case $A^* f = 0$ by Proposition 5.9. It follows that $A^* f$ is an alternating tensor on V. $\square$

Alternating tensors are very useful as we will come to see but one particular result is worth showing immediately. Suppose we have some linear transformation $B : V \to V$ mapping from a vector space V into itself. Then the restriction of the dual transformation to alternating n-tensors in $\mathcal{L}^n(V)$ given by $B^* : \mathcal{A}^n(V) \to \mathcal{A}^n(V)$ will multiply a given alternating tensor by some real number. This real number turns out to be the **determinant** of the transformation.

Theorem 5.13 (The Determinant in terms of $f \in \mathcal{A}^n(V)$): Let V be a vector space of dimension n with a linear transformation $B : V \to V$ mapping V into itself. Then the restriction $B^* : \mathcal{A}^n(V) \to \mathcal{A}^n(V)$ of the dual transformation to alternating n-tensors on V multiplies $f \in \mathcal{A}^n(V)$ by the determinant of B as in the following equation

$$B^* f = (\det B) f.$$

Proof: Since $k = n$ in this case, we only have one ascending k-tuple of integers being the set $\{1, \dots, n\}$. As such there is only one non-zero elementary alternating tensor $\Psi^{(1,\dots,n)} \in \mathcal{A}^n(V)$. This means we only need to check that the equation is true for $f = \Psi^{(1,\dots,n)}$. Let $a_1, a_2, \dots, a_n$ be a basis for V, then

$$\begin{aligned} B^* \Psi^{(1,\dots,n)}(a_1, a_2, \dots, a_n) &= \Psi_{(1,\dots,n)}(B(a_1), \dots, B(a_n)) \\ &= \sum_{\sigma \in S_k} (\mathrm{sgn}\sigma) e^{(1,\dots,n)}((B(a_{\sigma(1)}), \dots, B(a_{\sigma(n)})). \end{aligned}$$

Then using the definition of e^I we write this as

$$= \sum_{\sigma} (\mathrm{sgn}\sigma)(B_{1,\sigma(1)} \cdots B_{n,\sigma(n)})$$

which is just the formula for $\det B$. □

Recall that we showed $\mathcal{L}^k(V)$ was a vector space of tensors, with the tensor product $(\otimes)$ and addition as defined above. But since we are especially interested in alternating tensors we need to find a new product since the tensor product of two alternating tensors is almost never alternating.

Theorem 5.14 (The Wedge Product): Let V be a vector space. There exists an operation "$\wedge$": $\mathcal{A}^k(V) \times \mathcal{A}^l(V) \to \mathcal{A}^{k+l}(V)$ mapping two alternating tensors $f^{(k)}, g^{(l)} \mapsto f \wedge g$ such that the following properties hold:

1. **(Associativity):** $f \wedge (g \wedge h) = (f \wedge g) \wedge h$.
2. **(Multilinearity):** If f and g are of the same order,
$$(f + \lambda g) \wedge h = f \wedge h + \lambda(g \wedge h)$$
$$h \wedge (f + \lambda g) = h \wedge f + \lambda(h \wedge g)$$
3. **(Anti-commutativity):** $g \wedge f = (-1)^{kl}(f \wedge g)$.
4. If $\{a_i\}$ is a basis for V with dual basis $\{e^i\}$ for V^*. If $I = (i_1 < \cdots < i_k)$ is an ascending k-tuple of integers from $\{1, \dots, n\}$ then
$$\Psi^I = e^{i_1} \wedge \cdots \wedge e^{i_k},$$
where Ψ^I is an elementary alternating tensor.

5. If $T : V \to W$ is a linear transformation and f, g are alternating tensors on W then,

$$T^*(f \wedge g) = T^* f \wedge T^* g$$

The tensor $f \wedge g$ is called the **wedge product** of f and g uniquely defined by properties (1) to (4). The proof of this theorem is just as important as the result itself so we shall carefully walk through it step by step.

Proof Step (1): It is convenient for the proof to define the averaging operator $A : \mathcal{L}^k(V) \to \mathcal{L}^k(V)$. Let f be a k-tensor on V then A is defined by the equation

$$Af = \sum_{\sigma \in S_k} (\text{sgn } \sigma) f^\sigma.$$

Note that with A being defined in this way we can simply relate the basis elements of $\mathcal{L}^k(V)$ and $\mathcal{A}^k(V)$ as

$$\Psi^I = Ae^I$$

Claim (1): The transformation A has the following properties

1. A is linear.
2. Af is an alternating tensor.
3. If f is already an alternating tensor then $Af = (k!)f$.

Proof of Claim (1): The fact that A is linear comes from the fact that $f \mapsto f^\sigma$ is linear as proven in Lemma 5.8. To prove that Af is an alternating tensor, let $\tau \in S_k$ then

$$(Af)^\tau = \sum_{\sigma \in S_k} (\text{sgn } \sigma)(f^\sigma)^\tau = \sum_{\sigma \in S_k} (\text{sgn } \sigma) f^{\tau \circ \sigma}.$$

$$= (\text{sgn } \tau) \sum_{\sigma \in S_k} (\text{sgn } (\tau \circ \sigma)) f^{\tau \circ \sigma} = (\text{sgn } \tau) Af.$$

If f is already alternating, then $f^\sigma = (\text{sgn } \sigma) f$. It follows that

$$Af = \sum_{\sigma \in S_k} (\text{sgn } \sigma)^2 (f) = (k!) f$$

Which proves our claim. □

Proof Step (2): With the averaging operator we can turn any k-tensor into an alternating one and thus we can define an operation closed under alternating-ness as an extension of the tensor product. For $f \in \mathcal{A}^k(V)$ and $g \in \mathcal{A}^l(V)$ we define the **wedge product** of f and g by the equation

$$f \wedge g = \frac{1}{k!l!} A(f \otimes g).$$

We can use this definition to prove the defining properties and show that this wedge product is the one we want. The coefficient of $1/k!l!$ is mysterious but it is actually required for the associativity since there are $k!l!$ equivalent terms in the sum over permutations when f, g are alternating. Associativity is the hardest property to prove so we'll postpone it until later steps.

Proof Step (3): The multilinearity of the wedge product follows from the linearity of A and the fact that the tensor product is distributive, homogeneous and bilinear. We can prove anti-commutativity by computation.

Let f be a k-tensor and let g be an l-tensor (not necessarily alternating). Let π be the permutation which swaps the first k integers with the last l integers in the sequence $(1, \ldots, k+l)$, i.e. $(\pi(1), \ldots, \pi(k+l)) = (k+1, \ldots, k+l, 1, 2, \ldots, k)$. Then counting inversions, we see that $\text{sgn}\ \pi = (-1)^{kl}$. With this permutation we see that $(g \otimes f)^\pi = (f \otimes g)$ since

$$(g \otimes f)^\pi(v_1, \ldots, v_{k+l}) = g(v_{k+1}, \ldots, v_{k+l}) \cdot f(v_1, \ldots, v_k) = (f \otimes g)(v_1, \ldots, v_{k+l}).$$

We then compute

$$A(f \otimes g) = \sum_\sigma (\text{sgn}\ \sigma)(f \otimes g)^\sigma = \sum_\sigma (\text{sgn}\ \sigma)((g \otimes f)^\pi)^\sigma$$

$$= (\text{sgn}\pi) \sum_\sigma (\text{sgn}\ (\sigma \circ \pi))(g \otimes f)^{\sigma \circ \pi} = (-1)^{kl} A(g \otimes f).$$

Which verifies anti-commutativity.

Proof Step (4): We now begin to verify associativity, which begins by showing that the following is true.

Claim (2): If f and g are tensors of order k and l respectively such that $Af = 0$. Then $A(f \otimes g) = 0$.

Proof of Claim (2): Lets single out and consider one term in the expression for $A(f \otimes g)$ for example the term

$$(\text{sgn}\sigma) f(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \cdot g(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}).$$

If we now group together all terms in the expression with the same l-tuple of elements in the argument of g:

$$(\text{sgn}\sigma) \Big[\sum_{\tau \in S_k} (\text{sgn}\tau) f(v_{\sigma(\tau(1))}, \ldots, v_{\sigma(\tau(k))}) \Big] \cdot g(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}) \tag{5.3}$$

we identify that the expression in brackets is nothing but $Af(v_{\sigma(1)}, \ldots, v_{\sigma(k)})$ which is 0 by assumption. Repeating this process for all l-tuples in the argument of g covers the entire sum and therefore $A(f \otimes g) = 0$. $\square$

Proof Step (5): We now prove another claim in the direction of our result.

Claim (3): Let f be an arbitrary tensor and let h be an alternating m-tensor. Then we can show that

$$(Af) \wedge h = \frac{1}{m!} A(f \otimes h).$$

Proof of Claim (3): Let f have order k. Then the desired result can be written as

$$\frac{1}{k!m!} A((Af) \otimes h) = \frac{1}{m!} A(f \otimes h).$$

Then since the averaging operator is linear and $\otimes$ is distributive we can manipulate this as

$$A[(Af) \otimes h - (k!)(f \otimes h)] = 0 \implies A[(Af - k!f) \otimes h] = 0.$$

By step 4 we can reduce this to showing that $A[(Af - k!f)] = 0$ but since Af is already an alternating k-tensor $A(Af) = (k!)Af$ and the result follows. $\square$

Proof Step (6): We must show one final claim now before finally reaching associativity.

Claim (4): Let f, g, h be alternating tensors of degree k, l, m respectively. Then

$$(f \wedge g) \wedge h = \frac{1}{k!l!m!} A((f \otimes g) \otimes h).$$

Proof of Claim (4): By the definition of the wedge product, we can write

$$(f \wedge g) \wedge h = \frac{1}{k!l!} A(f \otimes g) \wedge h$$

But by the claim proven in step (5) this becomes

$$= \frac{1}{k!l!m!} A((f \otimes g) \otimes h)$$

Which proves our claim once again. $\square$

Proof Step (7): Using f, g, h as in the previous step, then

$$(k!l!m!)(f \wedge g) \wedge h = A((f \otimes g) \otimes h) = A(f \otimes (g \otimes h))$$

by the associativity of the tensor product. Now applying anti-commutativity from step (3) to permute the tensor product

$$= (-1)^{k(l+m)} A((g \otimes h) \otimes f) \underset{\text{Step (6)}}{=} (-1)^{k(l+m)} (k!l!m!)(g \wedge h) \wedge f$$

which by anti-commutativity becomes

$$(k!l!m!) f \wedge (g \wedge h).$$

Proof Step (8): Now we verify property (4) by proving a slightly more general result.

Claim (5): For any collection $f_1, \dots, f_k$ of 1-tensors,

$$A(f_1 \otimes \dots \otimes f_k) = f_1 \wedge \dots \wedge f_k$$

Proof of Claim (5): We prove this result by induction, the result is trivial if we have $k = 1$ 1-tensor. Assume the result is true for $k-1$ we then prove it for k. Let $F = (f_1 \otimes \dots \otimes f_{k-1})$. Then

$$A(F \otimes f_k) \underset{\text{Step (6)}}{=} (1!)(AF) \wedge f_k = (f_1 \wedge \dots \wedge f_{k-1}) \wedge f_k$$

by the induction hypothesis. Property (4) is an immediate result of this since $\Psi_I = Ae^I = A(e^{i_1} \otimes \dots \otimes e^{i_k})$. □

Proof Step (9): We now verify that the definition of the tensor product we have constructed is the unique representation satisfying properties (1) - (4). Given alternating tensors f and g of degree k and l respectively, we can write them uniquely in terms of elementary alternating tensors as

$$f = \sum_{\{I\}} b_I \Psi^I \qquad g = \sum_{\{J\}} c_J \Psi^J.$$

Where $\{I\}$ and $\{J\}$ are the sets of all ascending k-tuples and l-tuples of integers from the set $\{1, \dots, n\}$ respectively. By multilinearity, the tensor product of f and g in this representation is

$$f \wedge g = \sum_{\{I\}} \sum_{\{J\}} b_I c_J \Psi^I \wedge \Psi^J.$$

This shows that to compute the tensor product of any two tensors we only need to know how to compute wedge products

$$\Psi^I \wedge \Psi^J = (e^{i_1} \wedge \dots \wedge e^{i_k}) \wedge (e^{j_1} \wedge \dots \wedge e^{j_l}).$$

To do this we only need to know the following rules for wedge products between elementary tensors

$$\boxed{e^i \wedge e^j = -e^j \wedge e^i} \qquad \text{and} \qquad \boxed{e^i \wedge e^i = 0}$$

which follow from anti-commutativity. Therefore $\Psi^I \wedge \Psi^J$ has two possible values for any given pair of k and l tuples

$$\Psi^I \wedge \Psi^J = \begin{cases} 0 & \text{if any indices are the same,} \\ (\text{sgn}\pi)\Psi^K & \text{if otherwise.} \end{cases}$$

To be more precise, $\Psi^K K$ is the elementary alternating $k+l$-tensor with index K representing the $(k+l)$-tuple (I,J) after it has been rearranged into ascending order by the permutation π.

Proof Step (10): To verify the final property (5), let $T : V \to W$ be a linear transformation and f be an arbitrary tensor on W. One can convince themselves that $T^*(f^\sigma) = (T^* f)^\sigma$ since T^* is linear. As such, it follows that $T * (Af) = A(T^* f)$.

Now using this fact, let f and g be alternating tensors on W of degree k and l respectively. We compute

$$T^*(f \wedge g) = \frac{1}{k!l!} T^*(A(f \otimes g)) = \frac{1}{k!l!} A(T^*(f \otimes g)).$$

By Theorem 5.5 we know the dual transformation is distributive across the tensor product we obtain

$$\frac{1}{k!l!} A((T^* f) \otimes (T^* g)) = (T^* f) \wedge (T^* g).$$

Which completes the entire proof... (Yikes). □ □ □ □ □

This concludes the section on multilinear algebra we now have all we need to get back to manifolds and geometry.

6 Differential Forms

Definition (Differential Forms): Let M be a smooth manifold with boundary. A **differential form of order** k (or a k form) on M is a smooth function

$$\omega : \{(p, v_1, \dots, v_k) \in M \times \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{\text{k times}} \mid v_i \in T_p M\} \to \mathbb{R}$$

such that for every point p on the manifold, the restriction $\omega_p = \omega(p, \dots) : (T_p M)^k \to \mathbb{R}$ is an alternating k-tensor on $T_p M$. In other words, a k-form ω on M assigns to each point of M an element of $\mathcal{A}^k(T_p M)$. We group all such k-forms of a manifold into one set denoted by

$$\Omega^k(M) = \{\omega \mid \omega \text{ a smooth } k\text{-form on } M\}.$$

Remark: With the above definition of differential forms on a manifold, we note that $\Omega^0(M)$ is the set of all smooth real valued functions $C^\infty(M) : M \to \mathbb{R}$. In this case we see that $f \in \Omega^0(M)$ can be referred to as a 0-form on M. It is also common to refer to these functions as **scalar fields** in M. In addition, we let $\Omega^k_\delta(M)$ denote the set of not necessarily smooth k-forms on M.

Lemma 6.1 ($\Omega^k(M)$ **a Vector Space):** Let M be a smooth manifold with boundary. Then $\Omega^k(M)$ forms a vector space under pointwise addition and multiplication. Proof not required.

(Abuse of) Notation: In the definition that follows we will subtly denote the *i-th projection function* $\pi^i : \mathbb{R}^n \to \mathbb{R}$ which maps $(x_1, x_2, \ldots, x_n) \mapsto x_i$ by the coordinate with an upper index. So $x^i \equiv \pi^i$. We will soon resolve the ambiguity by showing that they are equivalent in the way that we use them. But in what follows we will have $dx^i = d(\pi^i(x_1, \ldots, x_n))$.

Definition (Elementary 1-Forms): Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$. Then for a point $p \in \mathbb{R}^n$, the set $\{(p, e_i)\}$ is called the **usual basis** for $T_x(\mathbb{R}^n)$. We define a 1-form dx^i on $\mathbb{R}^n$ by the equation

$$dx^i(p, e_j) = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases}$$

The forms $dx^1, dx^2, \ldots, dx^n$ are called **elementary** 1**-forms** on $\mathbb{R}^n$. Similarly, let $I = (i_1 < \cdots < i_k)$ be an ascending k-tuple of integers from the set $\{1, \ldots, n\}$. We define a k-form ψ^I on $\mathbb{R}^n$ by the equation

$$\psi^I = (dx^{i_1} \wedge \cdots \wedge dx^{i_k}) := dx^I.$$

We refer to ψ^I as **elementary** k**-forms** on $\mathbb{R}^n$. Note that for each $p \in \mathbb{R}^n$, the elementary 1-forms constitute the basis for $\mathcal{A}^1(T_p(\mathbb{R}^n))$ dual to the basis for $T_p(\mathbb{R}^n)$ and the corresponding $(k = n)$-tensor $\psi^I(p) \in \mathcal{A}^k(T_p(\mathbb{R}^n))$ is the elementary alternating tensor Ψ^I on $T_p(\mathbb{R}^n)$.

Both of these forms are manifestly smooth which can be seen from the following equations:

$$dx^i(p, v_1, \ldots, v_k) = v_i \qquad \text{and}$$

$$\psi^I(p)((p, \mathbf{v}_1), \ldots, (p, \mathbf{v}_k)) = \Psi^I(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \det[\mathbf{v}_1, \ldots, \mathbf{v}_k].$$

Definition (Induced Map): Let $f : M \to N$ be a smooth map between manifolds in $\mathbb{R}^n$. We define the transformation $f^* : \Omega^k(N) \to \Omega^k(M)$ by the following equation

$$(f^*\omega)_p(v_1, \ldots, v_k) = \omega_{f(p)}(Df(p)v_1, \ldots, Df(p)v_k) = (Df(p))^* \omega_{f(p)}(v_1, \ldots, v_k).$$

Here $p \in M$, $\omega \in \Omega^k(M)$ and $(v_1, \ldots, v_k) \in (T_pM)^k$.

Proposition 6.2 (The Induced Map is Well-Defined and Linear): If $f : M \to N$ be a smooth map between manifolds in $\mathbb{R}^n$, then the induced map $f^* : \Omega^k(N) \to \Omega^k(M)$ is well defined and linear.

Proof: The expression is well defined as a map between $(T_pM)^k$ and $(T_{f(p)}N)^k$ since each $Df(p) \cdot v_i \in T_f(p)N$. Linearity is also clear since $Df(p)$ is a linear transformation, we can show this is true also by computation, let $\omega, \eta \in \Omega^k(M)$ and $(v_1, \ldots, v_k) \in (T_{f(p)}N)^k$, then

$$(f^*(\omega + \eta))_p(v_1, \ldots, v_k) = (\omega + \eta)_{f(p)}(Df(p)v_1, \ldots, Df(p)v_k).$$

By Lemma 6.1 we can split this pointwise addition into separate terms as

$$= \omega_{f(p)}(Df(p)v_1, \ldots, Df(p)v_k) + \eta_{f(p)}(Df(p)v_1, \ldots, Df(p)v_k)$$

$$= [(f^*\omega)_p + (f^*\eta)_p](v_1, \ldots, v_k).$$

It remains to show that $f^*\omega$ is smooth. One can notice that $f^*\omega$ can be written as the composition of firstly a map between $T^{(k)}M$ and $T^{(k)}N$ which are the tangent bundles of M and N extended to k-tuples of vectors from T_pM and $T_{f(p)}N$ respectively. And the second map in the composition being the k-form $\omega : T^{(k)}N \to \mathbb{R}$. We already have that ω is smooth by definition, thus it suffices to show that the map $f_*^{(k)} : T^{(k)}M \to T^{(k)}N$ is smooth. Let $M \subset \mathbb{R}^n$ and $N \subset \mathbb{R}^m$. Let $\tilde{f}$ be a smooth extension of f, then $f_*^{(k)}$ extends to $\tilde{f}_*^{(k)}$. But $\tilde{f}_*^{(k)} : U \supset M \times (\mathbb{R}^n)^k \to \mathbb{R}^m \times (\mathbb{R}^n)^k$ is a composition of the maps

$$(\tilde{f} \times D\tilde{f} \times \mathbb{I}_{(\mathbb{R}^n)^k}) : U \times (\mathbb{R}^n)^k \to \mathbb{R}^m \times \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m) \times (\mathbb{R}^n)^k$$

and

$$\mathbb{I}_{\mathbb{R}^m} \times e_v^k : \mathbb{R}^m \times \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m) \times (\mathbb{R}^n)^k \to \mathbb{R}^m \times (\mathbb{R}^m)^k.$$

This means $\tilde{f}_*^{(k)}$ is a composition of smooth maps and hence smooth. Thus the restriction $f_*(k)$ to $T^{(k)}M$ is also smooth which proves our result for the induced map $f^*\omega$. □

Proposition 6.3 (Properties of the Induced Map): Let $g : M \to N$ and $f : N \to L$ be smooth maps between smooth manifolds with boundary, then

1. $\mathbb{I}_M^* = \mathbb{I}_{\Omega^k(M)}$.
2. $(f \circ g)^* = g^* \circ f^*$.

Proof: To prove property (1), we note that $\mathbb{I}_M$ is a smooth mapping from M onto itself and thus the transformation induced by $\mathbb{I}_M$ will be a smooth mapping from the space $\Omega^k(M)$ onto itself we simply apply the definition of the induced transformation for $\omega \in \Omega^k(M)$ to see

$$(\mathbb{I}_M^*\omega)_p(v_1, \ldots, v_k) = \omega_{\mathbb{I}_M(p)}(\mathbb{I}_{T_pM}v_1, \ldots, \mathbb{I}_{T_pM}v_k).$$

But $\mathbb{I}_M(p) = p$ and $\mathbb{I}_{T_pM}v_i = v_i$ therefore

$$(\mathbb{I}_M^*\omega)_p(v_1, \ldots, v_k) = \omega_p(v_1, \ldots, v_k) \Longleftrightarrow \mathbb{I}_M^* = \mathbb{I}_{\Omega^k(M)}.$$

Property (2) is simply obtained by applying the chain rule. Let $q = g(p)$ and $r = f(q)$. Then

$$((f \circ g)^*\omega)_p(v_1, \ldots, v_k) = \omega_{(f\circ g)(p)}(D(f \circ g)v_1, \ldots, D(f \circ g)v_k).$$

$$= \omega_{f(q)}((Df(q) \cdot Dg(p))v_1, \ldots, (Df(q) \cdot Dg(p))v_k) = (f^*\omega)_{g(p)}(Dg(p)v_1, \ldots, Dg(p)v_k). \tag{6.1}$$

$$= (g^*(f^*\omega))_p(v_1, \ldots, v_k) = ((g^* \circ f^*)\omega)_p(v_1, \ldots, v_k).$$

Which verifies our result. $\square$

Proposition 6.4 (Condition for Smoothness of k-forms): Let $M \subseteq \mathbb{R}^n$ be a smooth d-manifold. An arbitrary k-form $\omega \in \Omega_\delta^k(M)$ is smooth if for every point $p \in M$, there exists a **chart** $\alpha : U \to V$ such that $\alpha^*\omega \in \Omega_\delta^k(M)$ is smooth.

Proof: It suffices to show that ω is smooth at each $(p, v_1, \dots, v_k) \in T^{(k)}M$. First we observe that $T^{(k)}V$ is open in $T^{(k)}M$ which are respectively the tangent bundles of M and V extended to k-tuples of vectors from T_pM and T_pV, hence we only need to show that the restriction $\omega|_V : T^{(k)}V \to \mathbb{R}$ is smooth. But since $\alpha(U) = V$ and $\alpha^{-1}(V) = U$, we can write this restriction as the induced map of the composition

$$\omega|_V = ((\alpha^{-1*} \circ \alpha^*)\omega)|_V = (\alpha^{-1*}(\alpha^*\omega))|_V.$$

But $\alpha^*\omega$ is smooth by assumption, and α^{-1} is smooth since α is a coordinate patch. Thus $\omega|_V$ is smooth as a composition of smooth functions. $\square$

Proposition 6.5 (Further Condition for Smoothness of k-forms): If ω is a k-form defined on an open set $U \in \mathbb{R}^n$, then the k-tensor $\omega_p \in \mathcal{A}^k(\mathbb{R}^d)$ can be written uniquely in the form

$$\omega_p = \sum_{\{I\}} \omega_{p,I} \Psi^I \tag{6.2}$$

for some scalar function components $\omega_{p,I}$. Then ω is smooth iff the components $p \mapsto \omega_{p,I}$ are smooth for all ascending k-tuples I of integers from the set $\{1, \dots, n\}$.

Proof: That ω_p can be written uniquely in the form shown above follows from the uniqueness of the basis $\Psi^I \in \mathcal{A}^k(\mathbb{R}^d)$ of elementary alternating tensors on $\mathbb{R}^d$. To prove the more juicy statement here let an arbitrary k-form $\omega \in \Omega_\delta^k(\mathbb{R}^d)$ be given. We can express it in terms of elementary forms by the equation

$$\omega = \sum_{\{I\}} \omega_I \ \psi^I$$

The functions ψ^I are smooth, so ω is smooth if the component functions ω_I are smooth. Conversely , if ω is a smooth function of $(p, v_1, \dots, v_k) \in \mathbb{R}^d \times (T_p\mathbb{R}^d)^k$, then given an ascending k-tuple $J = (j_1, \dots, j_k)$ from $\{1, \dots, n\}$, the function

$$\omega_p(e_{j_1}, \dots, e_{j_k})$$

is smooth on the usual basis of $T_p\mathbb{R}^d$ but this is nothing but $\omega_{p,J}$. $\square$

Definition (Wedge Product of k-forms): Let $\omega \in \Omega^k(M)$, $\eta \in \Omega^l(M)$ then we define the wedge product of ω and η to be the $(k+l)$-form $\omega \wedge \eta \in \Omega^{k+l}(M)$ defined by the equation

$$(\omega \wedge \eta)_p(v_1, \dots, v_{k+l}) = (\omega_p \wedge \eta_p)(v_1, \dots, v_{k+l}).$$

We must make the following claim to complete this definition as it follows non-trivially.

Claim: The $(k+l)$-form $\omega \wedge \eta$ is smooth. Using the result of Proposition 6.5, we can assume that for $M \subseteq \mathbb{R}^d$ open, we can write the forms uniquely as

$$\omega = \sum_{\{I\}} \omega_I \ \psi^I \quad \text{and} \quad \eta = \sum_{\{J\}} \eta_J \ \psi^J$$

where I, J are ascending k and l-tuples of integers from the set $\{1, \dots, n\}$. Then

$$\omega \wedge \eta = \sum_{\{I\}} \sum_{\{J\}} (\omega_I \cdot \eta_J)(dx^I \wedge dx^J).$$

Just as in step (9) of the proof of Theorem 5.14, the wedge product $dx^I \wedge dx^J = dx^K$ where dx^K is either 0 if any two indices are the same or dx^K if $K = I \cup J$ is ascending. Clearly the wedge product of two elementary k-forms is then smooth and hence the wedge product of two forms is smooth. □

Proposition 6.5 (Induced Map on Wedge Products): Let $f : M \to N$ be a smooth function between open sets $M \subseteq \mathbb{R}^m$ and $N \subseteq \mathbb{R}^n$. Then for arbitrary forms ω, η on M, the induced map f^* is distributive across the wedge product

$$f^*(\omega \wedge \eta) = f^*\omega \wedge f^*\eta. \tag{6.3}$$

Proof: We show this by computation, let $p \in M$

$$f^*(\omega \wedge \eta)(p) = (Df(p))^*(\omega \wedge \eta)_{f(p)} \tag{6.4}$$

which by the definition of the wedge product of forms we can write as

$$(Df(p))^*(\omega_{f(p)} \wedge \eta_{f(p)}) = (Df(p))^*\omega_{f(p)} \wedge (Df(p))^*\eta_{f(p)})$$

since the induced transformation is linear. Then

$$= (f^*\omega)_p \wedge (f^*\eta)_p = (f^*\omega \wedge f^*\eta)_p.$$

□

The results up to this point in the section of differential forms, constitutes the formal algebraic background for how we want to proceed. In what follows, we will define the *differential* or *exterior derivative* which is an operator whose action is to convert k forms to $(k+1)$-forms. It will be denoted by the Roman letter d and as you have seen I have sneaked it into some of our definitions so far rather ambiguously but the following lemma should do away with any confusion.

Lemma 6.6 (Elementary One Forms in Terms of π^i): Let M be an open set in $\mathbb{R}^n$ and let $\pi^i : M \to \mathbb{R}$ be the i-th projection function defined by

$$\pi^i(x^1, \dots, x^n) = x^i$$

then we define the elementary 1-forms as $dx^i = d\pi^i$ where $d\pi^i$ is the (yet to be properly defined) differential of π^i.

Proof: Since π^i is a smooth function, $d\pi^i$ is a smooth 1-form and for $p \in M$, $v \in T_pM$ we compute

$$d\pi^i(p)(p,v) = D\pi^i(p) \cdot v = v_i,$$

which satisfies the definition of the elementary 1-form on M. Thus for I an increasing k-tuple any elementary k-form can be written as

$$dx^{i_1} \wedge \cdots \wedge dx^{i_k} = d(\pi^{i_1}) \wedge \cdots \wedge d(\pi^{i_k}) \iff dx^I = d(\pi^I).$$

$\square$

Definition (Differential of a 0-form): Let M be an open set in $\mathbb{R}^n$ and let $f : M \to \mathbb{R}$ be a smooth function. Then $f \in \Omega^0(M) = C^\infty(M)$ is a 0-form on M and we define the 1-form $df \in \Omega^1(M)$ by the equation $df = f^*dx$. This one form is called the **differential of f.**

We can unpack this definition a little bit further by considering the tensor assigned by df to a point $p \in M$. For this we consider the equation

$$df(p,v) \equiv df_p(v) = (f^*dx_p)(v) = Df(p) \cdot v.$$

Where one recalls that $dx \equiv dx^1 \in \Omega^1(\mathbb{R})$ is the elementary 1-form on the real numbers defined by $dx(p,v) = e_1(v) = v$. In conclusion, the differential of a 0-form is $df = Df$ and $Df(p) : T_pM \to T_{f(p)}\mathbb{R}$ is a $1 \times n$ matrix.

Proposition 6.7 (Coordinate Definition of df): Let $f \in C^\infty(M)$ where M is an open subset of $\mathbb{R}^d$. Then the basis of $\Omega^1(\mathbb{R}^d)$ in this space is the set of elementary 1-forms

$$\{dx^1, \ldots, dx^d\} \subseteq \Omega^1(\mathbb{R}^d).$$

Then the differential of f can be written uniquely in coordinates as

$$df = \sum_{i=1}^{d} \frac{\partial f}{\partial x^i} \cdot dx^i.$$

Proof: Let $p \in M$ and $v \in T_pM$, then we compute

$$df_p(v) = Df(p) \cdot v = \sum_{i=1} \frac{\partial f}{\partial x^i}(p) v^i = \sum_{i=1} \frac{\partial f}{\partial x^i}(p) dx^i_p(v).$$

$\square$

Proposition 6.8 (Induced Map on a Differential): Let $F : M \to N$ be a smooth map between smooth manifolds with boundary. Then

$$F^*df = d(F^*f) = d(f \circ F)$$

Proof: We prove with a simple computation

$$F^* df = F^*(f^* dx)$$

then applying Proposition 6.3,

$$= (f \circ F)^* dx = d(f \circ F).$$

$\square$

Theorem 6.9 (Linearity of df): Let M be an open set of $\mathbb{R}^d$. The operator $d : C^\infty(M) \to \Omega^1(M)$ is linear on 0-forms.

Proof: Let $f, g : M \to \mathbb{R}$ be 0-forms on M. Define $F : \mathbb{R}^2 \to \mathbb{R}$ by $F(x, y) = \lambda x + y$ where λ is a scalar. Then $dF = \lambda dx + dy$. Now define $f \times g : M \to \mathbb{R}^2$ by the mapping $x \mapsto (f(x), g(x))$. Then

$$d(\lambda f + g) = d(F \circ f \times g) = (f \times g)^* dF = (f \times g)^*(\lambda dx + dy).$$

Then by the linearity of the induced map this becomes

$$\lambda(f \times g)^* dx + (f \times g)^* dy = \lambda df + dg.$$

$\square$

Proposition 6.10 (Leibniz Rule for d): Let M be an open set of $\mathbb{R}^d$ and let Let $f, g : M \to \mathbb{R}$ be 0-forms on M. Then

$$d(f \cdot g) = f dg + g df$$

Proof: Let $F : \mathbb{R}^2 \to \mathbb{R}$ be defined by the mapping $(x, y) \mapsto xy$ such that $dF = xdy + ydx$. Let $f \times g$ be defined as in the preceding theorem, then

$$d(f \cdot g) = d(F \circ (f \times g)) = (f \times g)^* dF = (f \times g)^*(xdy + ydx)$$

by the linearity property of the induced map this becomes

$$(f \times g)^* x \cdot (f \times g)^* dy + (f \times g)^* y \cdot (f \times g)^* dx = f dg + g df$$

which follows by applying the induced map on x, y, dx, dy. $\square$

Now we try to prove the existence of the d operator for forms of arbitrary order. The proof of this theorem is again just as important as the result itself so we will take care to walk through it step by step.

Theorem 6.11 (The Exterior Derivative / Differential): Let $M = U$ be an open set in $\mathbb{R}^n$. There exists a unique linear transformation

$$d : \Omega^k(M) \to \Omega^{k+1}(M)$$

defined for $k \geq 0$ such that:

1. **(On 0-forms):** If $f \in C^\infty(M)$ a 0-form, then df is the 1-form
$$df_p(v) = Df(p) \cdot v.$$
2. **(Leibniz Rule)**If ω and η are forms of order k and l respectively, then
$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta.$$
3. For every form ω $d(d\omega) = 0$.
4. **(Naturality):** If $F : M \to N$ is any smooth map between smooth manifolds, then for $\omega \in \Omega^k(N)$
$$d(F^* \omega) = F^*(d\omega).$$

We call d the **differential operator**, and we call $d\omega$ the differential of ω.

Proof: We start by proving this theorem for the special case $M \subset \mathbb{H}^n$.
Proof Step (1): We verify that such a transformation is unique. First we show the following claim.

Claim (1): Conditions (2) and (3) on d imply that for any forms $\omega_1, \ldots, \omega_k$, we have
$$d(d\omega_1 \wedge \cdots \wedge d\omega_k) = 0$$

Proof of Claim (1): We prove by induction. If $k = 1$, this equation is simply a consequence of property (3). Suppose it is true for $k - 1$, then set $\eta = (d\omega_1 \wedge \cdots \wedge d\omega_k)$ and by the Leibniz rule we compute
$$d(d\omega_1 \wedge \eta) = d(d\omega_1) \wedge \eta \pm d\omega_1 \wedge d\eta$$
The first term on the RHS vanishes by property (3) and the second vanishes by the induction hypothesis. □

Claim (2): For any form ω, the differential of ω, $d\omega$ is entirely and uniquely determined by the value of d on 0-forms which is specified in property (1).

Proof of Claim (2): Let $\omega \in \Omega^k(M)$, then we can write ω uniquely in the form
$$\omega = \sum_I \omega_I dx^I.$$
Computing the differential, since d is linear we can bring it inside the sum:
$$d\omega = \sum_I (d(\omega_I dx^I)) = \sum_I d\omega_I \wedge dx^I + \omega_I d(dx^I)$$

where we have used the Leibniz rule. But from the result of claim (1) and since dx^I is a k-form on M, we know that $d(dx^I) = 0$. Then

$$d\omega = \sum_I d\omega_I \wedge dx^I$$

which shows that $d\omega$ is determined by the value of d on the 0-form components ω_I when ω is written uniquely as a linear combination of elementary k-forms dx^I. □

Proof Step (2): Claim (2) readily provides a definition of d for forms of order $k \geq 1$ so we can make the definition now and proceed to verify that it satisfies properties (1) - (3).

Definition (Differential of a k-form): Let $\omega \in \Omega^k(M)$ where M is an open set in $\mathbb{R}^n$, then the differential of ω is

$$d\omega = \sum_{\{I\}} d\omega_I \wedge dx^I$$

where $\{I\}$ is the set of all ascending k-tuples of integers from the set $\{1, \dots, n\}$.

Proof Step (3): This definition indeed satisfies properties (1) - (3) as we can check:

1. In the case that ω is a zero form, $I = \varnothing$ and $d\omega_p(v) = (\omega^* dx_p)(v) = D\omega(p) \cdot v$.

2. Both sides of the Leibniz rule equation for property (2) are linear in both ω and η so we don't need to consider the whole sum in the definition but just one term to check that the rule holds. Therefore let I and J be ascending k and l tuples of integers from the set $\{1, \dots, n\}$. We consider the k-form $\omega = \omega_I dx^I$ and $\eta_J dx^J$. If any two indices in I and J are the same the wedge product of the tensors vanishes and the proof holds. If $I \cap J = \varnothing$ then there exists a permutation σ which brings $I \cup J$ into an ascending $k + l$ tuple of integers with index K and the wedge product is determined by $dx^I \wedge (\mathrm{sgn}\sigma) dx^J = dx^K$. So we compute the differential of $(\omega \wedge \eta) \in \Omega^{k+l}(M)$ as

   $$d(\omega \wedge \eta) = d(\omega_I \eta_J (\mathrm{sgn}\sigma)(dx^M)).$$

   Then by claim (2) we evaluate this differential as

   $$= (\mathrm{sgn}\sigma)\omega_I d\eta_J \wedge dx^M + (\mathrm{sgn}\sigma)\eta_J d\omega_I \wedge dx^M$$

   $$= \omega_I d\eta_J \wedge (dx^I \wedge dx^J) + \eta_J d\omega_I \wedge (dx^I \wedge dx^J)$$

   $$= (-1)^k (\omega_I dx^I) \wedge (d\eta_J \wedge dx^J) + (d\omega_I \wedge dx^J) \wedge (\eta_J dx^J)$$

   where the sign $(-1)^k$ comes from the fact that dx^I is a k-form and $d\eta_J$ is a 1-form. Finally

   $$= (-1)^k \omega \wedge d\eta + d\omega \wedge \eta$$

   which proves the Leibniz rule for d.

3. By linearity, it suffices to consider $\omega = \omega_I dx^I$ then we compute

$$d(d(\omega_I dx^I)) = d(d\omega_I \wedge dx^I).$$

Then applying the Leibniz rule we have

$$d(d\omega_I) \wedge dx^I + \underbrace{d\omega_I \wedge d(dx^I)}_{=0 \ \text{as proven in claim (1)}} \implies d(d\omega) = d(d\omega_I) \wedge dx^I.$$

So we now return to coordinates to compute

$$d(d\omega_I) = d\left(\sum_{i=1}^{n} \frac{\partial \omega_I}{\partial x^i} dx^i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 \omega_I}{\partial x^j \partial x^i}(dx^j \wedge dx^i).$$

But we can rewrite this sum as

$$\sum_{i>j}\left(\frac{\partial^2 \omega_I}{\partial x^j \partial x^i} - \frac{\partial^2 \omega_I}{\partial x^i \partial x^j}\right)(dx^j \wedge dx^i) = 0$$

which follows from the equality of mixed partial derivatives.

This proves that there exists an operator d which satisfies properties (1) – (3) on $\omega \in \Omega^k(\mathbb{H}^n)$ uniquely for all $k \geq 0$. The extension to general M is barely different as can be seen in Theorem 30.4 of Munkres. We shall complete two more steps of this current proof to prove property (4) and then briefly proof the case of general $M \in \mathbb{R}^n$.

Proof Step (4): To prove naturality (which holds in general), let $F : U \to V$ be a smooth mapping with $U \subseteq \mathbb{R}^{d_1}$ and $V \subseteq \mathbb{R}^{d_2}$. We first prove the case where ω is a 0-form. Let y denote a general point of $\mathbb{R}^{d_2}$, then by computation

$$F^*(d\omega) = F^*\left(\sum_{i=1}^{d_2} \frac{\partial \omega}{\partial y^i} dy^i\right) = \sum_{i=1}^{d_2}((D_i\omega) \circ F) dF^i.$$

Computing the other side of the formula for $x \in \mathbb{R}^{d_1}$

$$d(F^*\omega) = d(\omega \circ F) = \sum_{j=1}^{d_1} D_j(\omega \circ F) dx^j.$$

We can apply the chain rule now by setting $y = F(x)$, giving

$$D(\omega \circ F)(x) = D\omega(y) \cdot DF(x)$$

but notice that $D\omega(y) \cdot$ the j-th column of $DF(x)$ is just $D_j(\omega \circ F)(x)$ thus

$$D_j(\omega \circ F)(x) = \sum_{i=1}^{d_2}(D_i\omega)(y) \cdot D_j F^i(x) \implies D_j(\omega \circ F) = \sum_{i=1}^{d_2}((D_i\omega) \circ F) \cdot D_j F^i.$$

Substituting this result into the original equation we have

$$d(F^*\omega) = \sum_{j=1}^{d_1}\sum_{i=1}^{d_2}((D_i\omega)\circ F)\cdot D_jF^i dx^j = \sum_{i=1}^{d_2}((D_i\omega)\circ F)dF^i.$$

As found for the LHS.

Now we prove the result for $k \geq 1$. Since F^* and d are linear it suffices as usual to only show this property for one particular I in the sum, i.e. Let $\omega = \omega_I dx^I$, then

$$F^*(d\omega) = F^*(d\omega_I \wedge dx^I) = F^*(d\omega_I) \wedge F^*(dx^I).$$

On the other hand

$$\begin{aligned} d(F^*\omega) &= d(F^*(\omega_I \wedge dx^I)) = d((F^*\omega_I) \wedge (F^*(dx^I))) \\ &= d(F^*\omega_I) \wedge F^*(dx^I) + (F^*\omega) \wedge 0. \end{aligned}$$

Comparing both computations the result follows. □

Proof for General $M \subseteq \mathbb{R}^n$: To show the existence of $d : \Omega^k(M) \to \Omega^{k+1}(M)$ satisfying properties (1)-(3). Let $\alpha : U \to V$ be a chart around $p \in M$. Then we define the differential of ω by the equation

$$(d\omega)_p := ((\alpha^{-1})^* d(\alpha^*\omega))_p$$

We claim that this definition is well defined independent of the choice of chart.

Proof: Suppose we have two coordinate patches α, β around p. By restricting their codomains we can assume that $\beta = \alpha \circ \varphi$, where $\varphi : U_1 \to U_2$ a diffeomorphism (we used this technique frequently in the sections on Manifolds). We then have that:

$$(\beta^{-1})^* d(\beta^*\omega) = (\varphi^{-1} \circ \alpha^{-1})^* d(\alpha \circ \varphi)^*\omega.$$

Then using properties of the induced map over a composition this becomes

$$= (\alpha^{-1})^*((\varphi^{-1})^* d(\varphi^* \circ (\alpha^*\omega)) = (\alpha^{-1})^* \underbrace{(\varphi^{-1})^*\varphi^*}_{= \text{ identity}} d(\alpha^*\omega).$$

Therefore d as defined indeed exists and can be readily constructed from charts. For naturality, the above definition implies that $d\alpha^*\omega = \alpha^* d\omega$ whenever α is a chart. We claim that this definition satisfies the defining properties of the differential operator. The proof of this fact follows from the fact that α^* commutes with all other operations and so the analysis is virtually the same as what we did for the upper half space but with a few extra letters along for the ride. We conclude the proof and hopefully the fact that we know have the entire algebra of differential forms and the corresponding operator d at our disposal may give us something to be proud of. □

7 Integration of Forms

Now that we have established what differential forms are, its time to reap the rewards and see how we can use them. The notation for the rest of the course can get a bit hand-wavy so I will be very clear around the ambiguous spots as we move. We begin with a review and extension of vector calculus.

Definition (Vector Field): A **vector field** is a smooth function $X : M \to TM$ such that for $p \in M$, $X(p) \in T_pM$. We define the set of all vector fields on a set M by the equation

$$\mathfrak{X}(M) = \{X : M \to TM \mid X \text{ a vector field}\}.$$

For example if $M \subset \mathbb{R}^d$ then we have a set of standard basis vectors e_i, $i = (1, \dots, d)$ and elements of $\mathfrak{X}(M)$ take the form

$$X(p) = \left(p\ ; \sum_{i=1}^{d} f_i(p)e_i\right) \quad \text{where} \quad f_i \in C^\infty(M).$$

Remark (Alternate Definition of Vector Fields): There is a natural map from the tangent bundle back to the original set $\pi : TM \to M$ defined by $\pi(p, v) = p \in M$. We can define a vector field on M as a **smooth section of the natural inclusion**, that is, a function $X : M \to TM$ such that $\pi \circ X = \mathbb{I}_M$ is a vector field on M.

For the rest of this section we will only consider subsets of Euclidean space such that $M \subset \mathbb{R}^d$ and $TM = M \times \mathbb{R}^d$. Let's see if we can relate the vector fields of a subset $M \subset \mathbb{R}^d$ to the forms on M we encountered in the previous section. The first thing to note is that both definitions are very similar in that a k-form is a linear combination of smooth functions with elementary 1-forms and elements of a vector field are linear combinations of smooth functions with standard basis vectors. In particular, k-forms on M roughly take the form

$$\Omega^0(M) = C^\infty(M)$$

$$\Omega^1(M) = \left\{\sum_{i=1}^{d} f_i dx^i \mid f_i \in C^\infty(M)\right\}$$

$$\vdots$$

$$\Omega^{d-1}(M) = \left\{\sum_{i=1}^{d} g_i dx^1 \wedge \dots \wedge dx^{i-1} \wedge \hat{dx^i} \wedge dx^{i+1} \wedge \dots \wedge dx^d \mid g_i \in C^\infty(M)\right\}$$

$$\Omega^d(M) = \{u dx^1 \wedge \dots \wedge dx^d \mid u \in C^\infty(M)\}.$$

Where $\hat{a}$ denotes the omission of the term a. To complete the connection between vector fields and forms we must define some particularly useful operations which translate scalar fields into vector fields and vice-versa.

Definition (Gradient, Divergence and Curl): Let M be an open subset of $\mathbb{R}^d$. Let $f : M \to \mathbb{R}$ be a **scalar field in** M (a smooth, real-valued function on M). We define a corresponding vector field in M called the **gradient** of f, by the equation

$$(\mathbf{grad}\ f)(p) = (p; D_1 f(p)e_1 + \cdots + D_d f(p)e_d).$$

Let $X \in \mathfrak{X}(M)$ be defined by the equation

$$X(p) = \left(p\ ; \sum_{i=1}^{d} g_i(p)e_i\right) \equiv (p\ ; g(p)),$$

where $g : M \to \mathbb{R}^d$. Then we define a corresponding scalar field in M called the **divergence of** X by the equation

$$(\mathbf{div}\ X)(p) = D_1 g_1(p) + \cdots + D_d g_d(p).$$

In the special case $d = 3$, we define a corresponding vector field to X in M called the **curl of** X by the equation

$$(\mathbf{curl}\ X)(p) = (p\,; (D_2 g_3 - D_3 g_2)e_1 + (D_3 g_1 - D_1 g_3)e_2 + (D_1 g_2 - D_2 g_1)e_3) = (p\,; \det[\mathbf{e}, \mathbf{D}, \mathbf{X}]).$$

The following theorem shows how these transformations relate vector and scalar fields to the exterior derivative on forms.

Theorem 7.1 ($\mathfrak{X}(M)$ and $\Omega^k(M)$ are isomorphic): Let M be an open subset of $\mathbb{R}^d$. Then there exist vector space isomorphisms h_i such that the each of the following diagrams commute respectively:

$$\begin{array}{ccc} C^\infty(M) & \xrightarrow{h_0} & \Omega^0(M) \\ \big\downarrow \mathbf{grad} & & \big\downarrow d \\ \mathfrak{X}(M) & \xrightarrow{h_1} & \Omega^1(M) \end{array}$$

$$\begin{array}{ccc} \mathfrak{X}(M) & \xrightarrow{h_{n-1}} & \Omega^{d-1}(M) \\ \big\downarrow \mathbf{div} & & \big\downarrow d \\ C^\infty(M) & \xrightarrow{h_d} & \Omega^d(M) \end{array}$$

i.e. $d \circ h_0 = h_1 \circ \mathbf{grad}$ and $d \circ h_{d-1} = h_d \circ \mathbf{div}$. Furthermore in the special case $d = 3$, the following diagram also commutes:

$$\begin{array}{ccc} \mathfrak{X}(M) & \xrightarrow{h_1} & \Omega^1(M) \\ \big\downarrow \mathbf{curl} & & \big\downarrow d \\ \mathfrak{X}(M) & \xrightarrow{h_2} & \Omega^d(M) \end{array}$$

i.e. $d \circ h_1 = h_2 \circ \mathrm{curl}$.

Proof: Let $g \in C^\infty(M)$ and $F \in \mathfrak{X}(M)$. We define the transformations h_i for $i = (1, \ldots, d)$ by the equations:

$$h_0 : C^\infty(M) \to \Omega^0(M) = \mathbb{I}_{C^\infty(M)}$$

$$h_1 : \mathfrak{X}(M) \to \Omega^1(M) \equiv h_1 F = \sum_{i=1}^d f_i dx^i$$

$$\vdots$$

$$h_{d-1} : \mathfrak{X}(M) \to \Omega^{d-1}(M) \equiv h_{d-1} F = \sum_{i=1}^d (-1)^{i-1} f_i dx^1 \wedge \cdots \wedge \hat{dx^i} \wedge \cdots \wedge dx^d$$

$$h_d : C^\infty(M) \to \Omega^d(M) \equiv h_d g = g dx^1 \wedge \cdots \wedge dx^d$$

Let's first verify the equation for the gradient using these definitions of h_0 and h_1. We compute

$$h_1(\mathbf{grad}\ g) = h_1\left(\sum_{i=1}^d D_i g\right) = \sum_{i=1}^d \frac{\partial g}{\partial x^i} dx^i = dg = (d \circ h_0) g$$

by definition. In a similar way we compute

$$(d \circ h_{d-1})(F) = d\left(\sum_{i=1}^d (-1)^{i-1} f_i dx^1 \wedge \cdots \wedge \hat{dx^i} \wedge \cdots \wedge dx^d\right)$$

$$= \sum_{i=1}^d \sum_{j=1}^d (-1)^{i-1} \frac{\partial f_i}{\partial x^j} dx^j \wedge dx^1 \wedge \cdots \wedge \hat{dx^i} \wedge \cdots \wedge dx^d.$$

The terms in the sum over j for which $i \neq j$ cancel since it would contain a wedge product of two identical elementary 1-forms so we are left with

$$(d \circ h_{d-1})(F) = \sum_{i=1}^d (-1)^{i-1} \frac{\partial f_i}{\partial x^i} dx^i \wedge dx^1 \wedge \cdots \wedge \hat{dx^i} \wedge \cdots \wedge dx^d$$

$$= \left(\sum_{i=1}^d \frac{\partial f_i}{\partial x^i}\right) dx^1 \wedge \cdots \wedge dx^d = (\mathbf{div}\ F) dx^1 \wedge \cdots \wedge dx^d = h_d(\mathbf{div}\ F).$$

In the special case $d = 3$ we can also show that $(d \circ h_1)F = (h_2 \circ \mathbf{curl})F$ by computing

$$dh_1(F) = d(f_1 dx^1 + f_2 dx^2 + f_3 dx^3) = D_2 f_1 dx^2 \wedge dx^1 + D_3 f_1 dx^3 \wedge dx^1 + D_1 f_2 dx^1 \wedge dx^2 + D_3 f_2 dx^3 \wedge dx^2$$

$$+ D_1 f_3 dx^1 \wedge dx^3 + D_2 f_3 dx^2 \wedge dx^3$$

$$= (D_2 g_3 - D_3 g_2) dx^2 \wedge dx^3 - (D_3 g_1 - D_1 g_3) dx^1 \wedge dx^3 + (D_1 g_2 - D_2 g_1) dx^1 \wedge dx^2$$

$$= h_2(\det[\mathbf{e}, \mathbf{D}, \mathbf{F}]) = (h_2 \circ \mathbf{curl}) F$$

□

Now we would like to learn why the differential forms we have constructed are the objects we hinted at in the discussion at the start of Section 5. In other words we want to know why and more importantly **how** we can integrate forms over manifolds in $\mathbb{R}^d$. So lets think back to our course on analysis in several real variables where we proved the following theorem.

Theorem 7.2 (Existence of Integration in $\mathbb{R}^d$): Let $dx^1dx^2\cdots dx^d$ be the infinitesimal volume element in $\mathbb{R}^d$, then there exists an operation called **integration** denoted by $\int_{\mathbb{R}^d} dx^1dx^2\cdots dx^d : L^1(\mathbb{R}^d) \to \mathbb{R}$ which satisfies the following properties:

$$\textbf{(1)}\ \text{(Fubini's Theorem)}\ \int_{\mathbb{R}^d} f(x^1,\ldots,x^d)\, dx^1\cdots dx^d = \int_{\mathbb{R}^{d-l}} \left[\int_{\mathbb{R}^l} f(x^1,\ldots,x^d) dx^1\cdots dx^l \right] dx^{l+1}\cdots dx^d.$$

$$\textbf{(2)}\ \text{(Fundamental Theorem of Calculus)}\ \int_{\mathbb{R}} F'(x)dx = F(\infty) - F(-\infty).$$

$$\textbf{(3)}\ \text{(Change of Variables)}\ \int_{\mathbb{R}^d} f(F(x))|\det DF| dx^1\cdots dx^d = \int_{\mathbb{R}^d} f(x) dx^1\cdots dx^d$$

where we have let $F : U_1 \to U_2$ be a diffeomorphism between open subsets of $\mathbb{H}^d$ such that Support$f \subseteq U_2$.

Okay so we know that such an operation exists in Euclidean space but our goal is to define $\int_M \omega$ where M is a d-manifold with boundary and $\omega \in \Omega^d(M)$. We can make an educated guess at how to define the integral using the procedure below in the special case that M is a compact manifold.

Idea: Suppose the manifold M is compact. We can cover M with finitely many coordinate patches $\alpha_i : U_i \to V_i$ and then choose a partition of unity $\{\varphi_i\}$ with each φ_i subordinate to V_i (vanishing outside of V_i). Then

$$\int_M \omega = \int_M \sum_i \varphi_i \omega = \sum_i \int_M (\varphi_i \omega)$$

Since the integrand vanishes outside of the sets V_i we can write it as

$$= \sum_i \int_{U_i} \alpha_i^*(\varphi_i \omega) = \sum_i \int_{U_i} f_i dx^i \cdots dx^d.$$

Where $f_i \in C^\infty(U_i)$ are the components of the d-forms $\alpha^*(\varphi_i\omega) \in \Omega^d(U_i)$. We want to set this as our definition in general but it must be independent of the choice of atlas $\{\alpha_i : U_i \to V_i\}$ which it is currently not. For example if $M = \mathbb{R}$ and the form we are trying to integrate is given by $\omega = f(x)dx$ where f is a strictly positive function, then the charts $\alpha(x) = x$ and $\alpha(x) = -x$ produce integrals which are off by a minus sign from each other so we need to define a condition on manifolds known as **orientability**.

Definition (Compact Support of a k-form): The k-form $\omega \in \Omega^k(M)$ is said to have **compact support** if

$$\textbf{Support}\ \omega = \overline{\{p \in M \mid \omega_p \neq 0\}}$$

is compact. From here we let $\Omega_c^k(M)$ denote the space of all k-forms in M with compact support.

In the special case that M is an open subset of the upper half space $\mathbb{H}^d$ we need not worry about orientability so we can make the definition stand.

Definition (Integration of Forms over $M \underset{\textbf{open}}{\subseteq} \mathbb{H}^d$): Let $M \subseteq \mathbb{H}^d$ be an open subset and $\omega \in \Omega_c^d(M)$ then as in the opening procedure we define

$$\int_M \omega = \int_{\mathbb{R}} u dx^1 \cdots dx^d \quad \text{where} \quad \omega = u dx^1 \wedge \cdots \wedge dx^d.$$

Definition (Orientation Preservation or Reversal of Diffeomorphsims): Let $F : M \to N$ be a diffeomorphism mapping between open subsets M and N of $\mathbb{H}^d$. Then for all $p \in M$, we call F **orientation preserving** if $\det[DF(p)] > 0$ and **orientation reversing** if $\det[DF(p)] < 0$.

We are able to make this definition for all $p \in M$ since F being a diffeomorphism means $\det[DF(p)]$ is constant and hence has constant sign on connected components.

Proposition 7.3 (Integration of forms under the induced map of an orientation preserving diffeomorphism): Let $F : M \to N$ be a diffeomorphism mapping between open subsets M and N of $\mathbb{H}^d$. Suppose also that F is orientation preserving. Then

$$\int_M F^* \omega = \int_N \omega \quad \text{for all} \quad \omega \in \Omega_c^d(N)$$

Proof: Consider ω as written in the form of elementary 1-forms $\omega = u dx^1 \wedge \cdots \wedge dx^d$. Lets first evaluate the transformation $F^*\omega$ at some point $p \in M$.

$$(F^*\omega)_p = (F^*u)_p \cdot DF(p)^* dx^1 \wedge \cdots \wedge dx^d = (u \circ F)(p) \cdot \det(DF(p)) \cdot dx^1 \wedge \cdots \wedge dx^d$$

Now the integral becomes

$$\int_M F^* \omega = \int_{\mathbb{R}^d} (u \circ F)(\det DF) dx^1 \cdots dx^d$$

Now since F is orientation preserving we know that $\det(DF) = |\det(DF)|$ and so we can directly apply the change of variables theorem to write this as

$$\int_M F^* \omega = \int_{\mathbb{R}^d} u dx^1 \cdots dx^d = \int_N \omega$$

Through this final step we can clearly see that if F was **orientation reversing** then the integrals would be equal up to a minus sign. □

Now we can jump back into manifolds and see what the analogue of the orientability condition is for charts which cover a given manifold. (It should be quite clear at this point that charts are going to play a very important role in our ability to integrate forms over manifolds.)

Definition (Orientations and Positive Charts): Let M be a manifold with boundary. Two charts $\alpha_{i,j} : U_{i,j} \to V_{i,j}$ are said to **overlap positively** if their transition function

$$\alpha_j^{-1} \circ \alpha_i : \alpha_i^{-1}(V_i \cap V_j) \to \alpha_j^{-1}(V_1 \cap V_2)$$

is orientation preserving. That is $\det[D(\alpha_j^{-1} \circ \alpha_i)] > 0$. If M can be covered by a collection of charts that pairwise positively overlap then we say that M is an **orientable manifold** and the choice of charts in this collection is called an **orientation** on M usually denoted $\{\alpha_i\}$.
From here on we can use the notation $(M, \ \{\alpha_i\})$ to denote an oriented manifold with boundary.

Lemma 7.4 (Orientation of Positive Patches is Maximal): Let $(M, \ \{\alpha_i\})$ be an oriented manifold with boundary. We say that a coordinate patch is **positive** if it overlaps positively with all α_i. Then the collection of all positive patches

$$\mathfrak{B} = \{\beta : U \to V \mid \beta \ \text{ a positive patch } \}.$$

is an orientation containing $\{\alpha_i\}$ and is not properly contained within any other orientation. In other words, $\mathfrak{B}$ is the maximal orientation of M.

Proof: We show that $\mathfrak{B}$ does in fact form an orientation, that is, we show that any two elements $\beta_i, \beta_j \in \mathfrak{B}$ satisfy $D(\beta_j^{-1} \circ \beta_i)(x) > 0$. To compute the transition map, we are free to take some $\alpha_k \in \{\alpha_i\}$, then write

$$\begin{aligned} \beta_j^{-1} \circ \beta_i &= \beta_j^{-1} \circ \alpha_k^{-1} \circ \alpha_k \circ \beta_i \\ &= (\alpha_k \circ \beta_j)^{-1} \circ (\alpha_k \circ \beta_i). \end{aligned}$$

Now

$$\begin{aligned} D(\beta_j^{-1} \circ \beta_i)(x) &= D((\alpha_k \circ \beta_j)^{-1} \circ (\alpha_k \circ \beta_i))(x) \\ &= D(\alpha_k \circ \beta_j^{-1})(\alpha_k \circ \beta_i(x)) D(\alpha_k \circ \beta_i)(x) \end{aligned}$$

by the chain rule. Now taking the determinant of this equation

$$\begin{aligned} \det[D(\beta_j^{-1} \circ \beta_i)(x)] &= \det[D(\alpha_k \circ \beta_j^{-1})(\alpha_k \circ \beta_i(x)) D(\alpha_k \circ \beta_i)(x)] \\ &= \det[D(\alpha_k \circ \beta_j^{-1})(\alpha_k \circ \beta_i(x))] \det[D(\alpha_k \circ \beta_i)(x)]. \end{aligned}$$

Now since we already know $\{\alpha_i\}$ is an orientation, this means that α_k overlaps positively with β_i, β_j by definition. Thus both determinants in the above equation are positive and hence $\det[(\beta_j^{-1} \circ \beta_i)(x)] > 0$ so elements of $\mathfrak{B}$ overlap positively and $\mathfrak{B}$ is an orientation.

Forming an orientation, any element $\alpha_k : U_k \to V_k$ of $\{\alpha_i\}$ overlaps positively with any other $\alpha_j \in \{\alpha_i\}$, furthermore each mapping in this collection maps between U and V by definition so $\alpha_k \in \mathfrak{B}$ for all k. Therefore $\{\alpha_i\} \subset \mathfrak{B}$. To show that $\mathfrak{B}$ is maximal in this respect, we must show that a chart κ is only in $\mathfrak{B}$ if it overlaps positively with every $\beta_i \in \mathfrak{B}$. We can show this by a similar computation:

$$\kappa^{-1} \circ \beta_i = \kappa^{-1} \circ \alpha_k \circ \alpha_k^{-1} \circ \beta_i$$

$$= (\kappa^{-1} \circ \alpha_k) \circ (\alpha_k^{-1} \circ \beta_i).$$

Now

$$D(\kappa^{-1} \circ \alpha_k)(x) = D((\kappa^{-1} \circ \alpha_k) \circ (\alpha_k^{-1} \circ \beta_i))(x)$$
$$= D(\kappa^{-1} \circ \alpha_k)(\alpha_k^{-1} \circ \beta_i(x)) D(\alpha_k^{-1} \circ \beta_i)(x)$$

by the chain rule. Now taking the determinant of this equation

$$\det[D(\kappa^{-1} \circ \beta_i)(x)] = \det[D(\kappa^{-1} \circ \alpha_k)(\alpha_k^{-1} \circ \beta_i(x)) D(\alpha_k^{-1} \circ \beta_i)(x)]$$

$$= \det[D(\kappa^{-1} \circ \alpha_k)(\alpha_k^{-1} \circ \beta_i(x))] \det[D(\alpha_k^{-1} \circ \beta_i)(x)].$$

We already know the second determinant is positive since $\{\alpha_i\} \subset \mathfrak{B}$, so we see that the following statement holds

$$\det[D(\kappa^{-1} \circ \beta_i)(x)] > 0 \iff \det[D(\kappa^{-1} \circ \alpha_k)] > 0.$$

In other words, a chart κ only overlaps positively with elements of $\mathfrak{B}$ if it overlaps positively with elements of $\{\alpha_i\}$. So we have that κ overlaps positively with all $\beta_i \in \mathfrak{B}$ if and only if it is in $\mathfrak{B}$ (since being in $\mathfrak{B}$ requires positively overlapping with all α_k). Hence $\mathfrak{B}$ is the maximal orientation of M. □

Definition (Orientation of a 0-Manifold): An orientation on a 0 -dimensional manifold (a discrete collection of points) is the choice of a function $\sigma : M \to \{\pm 1\}$.

Remark (The Tricky Case of 1-Manifolds with Boundary): There are a flurry of counter examples of 1-manifolds which are orientable when their boundary is excluded but don't seem to be orientable when it is. For example when a line becomes a line segment with end points then the charts at each end point induce an orientation pointing in opposite directions to each other and they thus overlap negatively. The problem here being that the charts have domain in $\mathbb{H}^1$. So we will make a special exception for 1 - manifolds with boundary since it turns out that each and every one of them is orientable:

Convention: In the case of a 1-manifold with boundary, we will allow the domains of the coordinate patches covering M to be open sets in the negative reals $(-\infty, 0]$ as well as the upper half space $\mathbb{H}^1$.

Lemma 7.5 (Composition of Orientations with Diffeomorphisms): Let M be a d-manifold with boundary and orientation $\{\alpha_i\}$ and let $\tau : \mathbb{H}^d \to \mathbb{H}^d$ be any diffeomorphism. Then $\{\alpha_i \circ \tau\}$ is also an orientation.

Proof: For notational convenience lets take some point $p \in M$, then for two charts $\alpha_j, \alpha_k \in \{\alpha_i\}$ around p we will have $\alpha_j(x) = p = \alpha_k(y)$ for some $x, y \in \mathbb{H}^d$. Now to see that $\{\alpha_i \circ \tau\}$ forms an orientation we must compute the determinant of the transition map between the two general elements $\alpha_j \circ \tau$ and $\alpha_k \circ \tau$ to show that they overlap positively. We compute

$$(\alpha_k \circ \tau)^{-1} \circ (\alpha_j \circ \tau)$$

Now we want to take the derivative of this expression at the point $\tau^{-1}(x)$ which is the preimage of the point p of the point with respect to the composition $\alpha_j \circ \tau$. So we proceed

$$D((\alpha_k \circ \tau)^{-1} \circ (\alpha_j \circ \tau))(\tau^{-1}(x))$$
$$= D((\alpha_k \circ \tau)^{-1}(\alpha_j \circ \tau(\tau^{-1}(x)))D(\alpha_j \circ \tau)(\tau^{-1}(x))$$

by the chain rule. Now

$$= D(\tau^{-1} \circ \alpha_k^{-1})(p)D(\alpha_j \circ \tau)(\tau^{-1}(x)),$$

and then we apply the chain rule again to obtain

$$D\tau^{-1}(\alpha_k^{-1}(p)) \cdot D\alpha_k^{-1}(p) \cdot D\alpha_j(\tau(\tau^{-1}(x))) \cdot D\tau(\tau^{-1}(x))$$

$$D\tau^{-1}(y) \cdot D\alpha_k^{-1}(p) \cdot D\alpha_j(x) \cdot D\tau(\tau^{-1}(x))$$

Now note that we can apply the reverse chain rule to the α derivatives since $D\alpha_k^{-1}(p) \cdot D\alpha_j(x) = D\alpha_k^{-1}(\alpha_j(x))D\alpha_j(x) = D(\alpha_k^{-1} \circ \alpha_j)(x)$. This leaves

$$D((\alpha_k \circ \tau)^{-1} \circ (\alpha_j \circ \tau))(\tau^{-1}(x)) = D\tau^{-1}(y)D(\alpha_k^{-1} \circ \alpha_j)(x)D\tau(\tau^{-1}(x)).$$

Now taking the determinant of the right hand side

$$\det[D\tau^{-1}(y)D(\alpha_k^{-1}\circ\alpha_j)(x)D\tau(\tau^{-1}(x))] = \det[D\tau^{-1}(y)]\det[D\tau(\tau^{-1}(x))]\cdot\det[D(\alpha_k^{-1}\circ\alpha_j)(x)]$$

where we have used that since τ is a diffeomorphism $\det[D\tau]\det[D\tau^{-1}] = 1$. We already know that the elements α_k, α_j overlap positively since they are in the collective orientation $\{\alpha_i\}$. This means that our equation is positive

$$\det[D((\alpha_k \circ \tau)^{-1} \circ (\alpha_j \circ \tau))(\tau^{-1}(x))] > 0$$

and $\alpha_j \circ \tau$ and $\alpha_k \circ \tau$ overlap positively for all $\alpha_k, \alpha_j \in \{\alpha_i\}$. Thus $\{\alpha_i \circ \tau\}$ is an orientation whenever $\{\alpha_i\}$ is an orientation. □

Lemma 7.6 (Positive Charts with Connected Domain): Let $(M, \{\alpha_i\})$ be an oriented manifold, let $\alpha : U \to V$ be a coordinate patch where U is connected. Suppose there exists $\alpha_i : U_i \to V_i$ and $x \in U$ such that $\alpha(x) \in V_i$ and that $\det[D(\alpha_i^{-1} \circ \alpha)(x)] > 0$. Then $\{\alpha_i\} \cup \{\alpha\}$ is an orientation.

Proof: We are essentially aiming to prove that if a chart α has connected domain U and overlaps positively with **one of** the charts in $\{\alpha_i\}$ then it overlaps positively with **all of** the charts in $\{\alpha_i\}$.

Define $\lambda : U \to \{\pm 1\}$ as follows. For $y \in U$ we let $\lambda(y) = \mathrm{sgn}(\det D(\alpha_j \circ \alpha)(y))$ for any positive chart $\alpha_j : U_j \to V_j$ such that $p = \alpha(y) \in V_j$ so that the composition makes sense. This is well defined since we can always insert the identity $\alpha_k \circ \alpha_k^{-1}$ in between the transition function where α_k is a positive patch and as shown in the previous two proofs the chain rule tells us that

$$\det[D(\alpha_j^{-1} \circ \alpha)(y)] = \det[D(\alpha_j^{-1} \circ \alpha_k)(p)] \cdot \det[D(\alpha_k^{-1} \circ \alpha)(y)]$$

where the middle term is positive since both charts are in the orientation $\{\alpha_i\}$. Regardless of whether the sign is positive or negative, from the above equation it is easy to see that λ is continuous or **locally constant** and since U is connected this means that it is constant anywhere. Finally by our assumption that α overlaps positively with at least one patch, the fact that we have now shown that if it overlaps positively with one then it does with any general α_k means that α overlaps positively with all patches in the orientation $\{\alpha_i\}$. $\square$

Corollary 7.7 (A Diffeomorphism which Reverses an Orientation): Let $(M, \{\alpha_i\})$ be an oriented manifold of dimension $d > 1$. Let $\alpha : U \to V$ be a coordinate patch where U is connected. Then either α or $\alpha \circ \tau$ is a positive patch. Here $\tau : \mathbb{H}^d \to \mathbb{H}^d$ is an orientation reversing diffeomorphism given by

$$\tau(x^1, x^2, \dots, x^d) = (-x^1, x^2, \dots, x^d)$$

Proof: We conduct first a computation to determine a condition for the composite $\alpha \circ \tau$ of the patch α and the orientation reversing diffeomorphism τ to overlap positively with patches in the orientation of M. Let $\alpha_j, \alpha_k \in \{\alpha_i\}$ be patches in the orientation and let $p \in \mathbb{H}^d$ such that $\tau(p) \in U$. Then

$$D(\alpha_j^{-1} \circ (\alpha \circ \tau))(p) = D(\alpha_j^{-1} \circ \alpha_k \circ \alpha_k^{-1} \circ \alpha \circ \tau)(p).$$

$$D((\alpha_k^{-1} \circ \alpha_j)^{-1} \circ (\alpha_k^{-1} \circ \alpha) \circ \tau)(p)$$

which by the chain rule becomes

$$D((\alpha_k^{-1} \circ \alpha_j)^{-1}(\alpha_k^{-1} \circ \alpha \circ \tau(p))D(\alpha_k^{-1} \circ \alpha)(\tau(p))D\tau(p).$$

Now taking the determinant of both sides of the equation we find

$$\det[D(\alpha_j^{-1} \circ (\alpha \circ \tau))(p)] = \det[D((\alpha_k^{-1} \circ \alpha_j)^{-1}(\alpha_k^{-1} \circ \alpha \circ \tau(p))] \det[D(\alpha_k^{-1} \circ \alpha)(\tau(p))] \det[D\tau(p)].$$

We already know that α_j and α_k overlap positively since they are in the orientation so the first determinant on the RHS is positive. Furthermore the final determinant is negative since τ is an orientation reversing diffeomorphism. Thus $\det[D(\alpha_j^{-1} \circ (\alpha \circ \tau))(p)]$ is only positive if $\det[D(\alpha_k^{-1} \circ \alpha)(\tau(p))] \det[D\tau(p)]$ is negative. Thus $\alpha \circ \tau$ is a positive patch if and only if α is a negative patch. The contrapositive must also be true since the LHS is only negative if α is positive so $\alpha \circ \tau$ is a negative patch if and only if α is a positive patch.

Hence we have found that either but not both α or $\alpha \circ \tau$ is a positive patch. $\square$

For $d = 1$ we have the same result by defining τ instead as the mapping $\tau : [0, \infty) \to (-\infty, 0]$ mapping $x \mapsto -x$.
Now recall that the boundary of a manifold is also a manifold, it becomes an important task to classify how orientability translates between both objects which can be achieved by the following proposition.

Proposition 7.8 (The Corresponding Manifold with Boundary of an Orientable Manifold Without Boundary is Orientable): Let M be a manifold with boundary. Suppose the interior $\mathring{M} = M - \partial M$ is orientable, then so is M. More precisely, if $\mathrm{A} = \{\alpha_i\}$ is an orientation on $\mathring{M}$, then there exists a unique orientation $\mathfrak{B}$ of M which dominates A.

Proof: We define a set of positively overlapping coordinate coordinate patches over the entire manifold M as

$$\mathfrak{B} = \{\beta : U \to V \mid \beta \text{ a patch on } M \text{ overlapping positively with all } \alpha_i\text{'s}\}$$

It is clear that $\mathrm{A} \subset \mathfrak{B}$ from this definition. That $\mathfrak{B}$ itself is an orientation follows from the fact that if $\beta : U \to V$ is a patch on M, then $\beta|_{U \cap \mathring{\mathbb{H}}^d}$ is a patch on $\mathring{M}$. So let $\beta_1, \beta_2 \in \mathfrak{B}$ be coordinate patches of the form $\beta_i : U_i \to V_i$ where the domains U_i are open subsets in the upper half space $\mathbb{H}^d$. WLOG we set $V_1 = V_2 \equiv V$, now to show that $\mathfrak{B}$ is an orientation on M we must show that

$$\Phi = \beta_1^{-1} \circ \beta_2 : U_2 \to U_1$$

is orientation preserving. But since we know that $\mathring{M}$ is orientable we know that the restriction

$$\Phi|_{U_2 \cap \mathring{\mathbb{H}}^d} = \beta_1|^{-1}_{U_1 \cap \mathring{\mathbb{H}}^d} \circ \beta_2|_{U_2 \cap \mathring{\mathbb{H}}^d}$$

is orientation preserving. Now for simplicity of notation lets define the continuous function

$$u : U_2 \to \mathbb{R} \quad \text{as} \quad u(p) = \det[D\Phi(p)].$$

Then we know that the image of this function satisfies $u(U_2) \subset \mathbb{R} - \{0\}$. And under the same restriction away from the boundary of $\mathbb{H}^d$ we know that $u(U_2 - \partial\mathbb{H}^d) = u(U_2 \cap \mathring{\mathbb{H}}^d) \subset \mathbb{R}^{>0}$. But then Φ must be orientation preserving since

$$u(U_2) = u(\overline{U_2 - \partial\mathbb{H}^d}) \subseteq \overline{u(U_2 - \partial\mathbb{H}^d)} \subseteq \mathbb{R}^{>0}.$$

Then to show that the patches in $\mathfrak{B}$ cover M, let $p \in M$ and choose a chart $\beta : U \to V$ around p. WLOG let U be a connected set. Then by Corollary 7.7 we know that either β or $\beta \circ \tau$ overlaps positively with some and hence any α_i in the orientation of $\mathring{M}$ where τ is the orientation reversing diffeomorphism. This shows that we can pick any point on the boundary and find a patch around that point which itself or its reversal is in $\mathfrak{B}$, this takes care of covering the boundary while the interior is already covered by $\{\alpha_i\}$ which are trivially in $\mathfrak{B}$. □

Corollary 7.9 (Orientation of Manifolds Obtained by Regular Values): If 0 is a regular value of the smooth function $f : \mathbb{R}^d \to \mathbb{R}$ then the manifold $f^{-1}([0, \infty)) = M$ carries a natural orientation.

In the following section we will make the connection we need between these orientations and differential forms so that we can be sure about the definition of integration we want to make. An orientation on a manifold M is essentially a partition of the bases of T_pM for every point $p \in M$ into "positive" and "negative" ones since there are choices for the direction of the tangent space at each point on the manifold. An orientation on M *is the choice we make about which set of basis vectors for T_pM qualify as positive or negative* i.e. a function which assigns a value $\{\pm 1\}$ to each $(p; v_1, v_2, \dots)$. But this type of function is exactly the same as what we saw during our study of differential forms which of course become alternating tensors when evaluated at a point. So we can begin to see that there are a special type of forms we can use which do not vanish over the region of interest (subordinate support) that simultaneously can be integrated and also represent the orientation of the manifold over which we would like to integrate. We will refer to such objects as volume forms in what follows.

Definition (Volume Forms): Let M be a d-manifold. Let $\omega \in \Omega^d(M)$ such that for all $p \in M$, the corresponding tensor $\omega_p \neq 0$. Then ω is called a **volume form**.

Proposition 7.10 (Induced Orientation of a Volume Form): Any volume form ω induces an orientation by defining a chart $\alpha : U \to V$ to be **positive** if

$$\alpha^*\omega = u dx^1 \wedge \cdots \wedge dx^d \quad \text{where} \quad u(x) > 0 \quad \forall x \in U$$

Proof: Let α_1, α_2 be two patches which satisfy the definition of **positive** as stated here. WLOG let $V_1 = V_2 \equiv V$. We need to show that the function

$$\Phi = \alpha_1^{-1} \circ \alpha_2 : U_2 \to U_1$$

is orientation preserving. Since both patches are positive we already know that for ω a volume form,

$$\alpha_1^*\omega = u_1 dx^1 \wedge \cdots \wedge dx^d \quad \text{and} \quad \alpha_2^*\omega = u_2 dx^1 \wedge \cdots \wedge dx^d$$

where u_1, u_2 are strictly positive functions. In a slight trick we now apply $(\alpha_1^{-1})^*$ on the left of the first equation to obtain

$$(\alpha_1^{-1})^*\alpha_1^*\omega = (\alpha_1^{-1})^* u_1 dx^1 \wedge \cdots \wedge dx^d \implies \omega = (\alpha_1^{-1})^* u_1 dx^1 \wedge \cdots \wedge dx^d.$$

For the second equation this has the consequence that

$$\alpha_2^*\omega = \alpha_2^*((\alpha_1^{-1})^* u_1 dx^1 \wedge \cdots \wedge dx^d) = (\alpha_1^{-1} \circ \alpha_2)^*(u_1 dx^1 \wedge \cdots \wedge dx^d) = \Phi^*(u_1 dx^1 \wedge \cdots \wedge dx^d)$$
$$= \underbrace{\Phi^* u_1}_{>0} \cdot (\det D\Phi) dx^1 \wedge \cdots \wedge dx^d = \underbrace{u_2}_{>0} dx^1 \wedge \cdots \wedge dx^d.$$

Hence α_1 and α_2 overlap positively if they satisfy the volume form definition of positive and in this respect all such patches overlap pairwise positively and we can always find a set of positive patches covering M by choosing charts around each point with U connected and choosing either α or $\alpha \circ \tau$ so that they overlap positively with all other charts. □

This proposition is key, we have now demanded that integration of volume forms over a manifold is non-negative and it turns out that this can be achieved for **any orientation**.

Proposition 7.11 (All Orientations have a Corresponding Volume Form): Any orientation $\{\alpha_i\}$ is an orientation induced by some volume form.

Proof: Let $\{\alpha_i : U_i \to V_i\}$ be an orientation on a d-manifold M. We choose a partition of unity $\{\varphi_i\}$ subordinate to V_i. Let the d-form ω be given by the following unique representation

$$\omega = \sum_i \varphi_i (\alpha_i^{-1})^* (dx^1 \wedge \cdots dx^d).$$

(this is the same construction of ω we obtained in the last proof but with the positive function u_1 now replaced by a partition of unity). Then for one, ω is smooth. We can also show that ω induces the orientation $\{\alpha_i\}$. Let $\beta : U \to V$ be positive with respect to $\{\alpha_i\}$, then

$$\beta^*\omega = \sum_i \underbrace{\beta^*\varphi_i}_{>0} \cdot \beta^*((\alpha_i^{-1})^* dx^1 \wedge \cdots \wedge dx^d)$$

$$= \sum_i \beta^*\varphi_i \cdot \underbrace{(\alpha_i^{-1} \circ \beta)^*}_{>0} dx^1 \wedge \cdots \wedge dx^d = \underbrace{\left(\sum_i \beta^*\varphi_i \cdot \det[D(\alpha_i^{-1} \circ \beta)] \right)}_{>0} dx^1 \wedge \cdots \wedge dx^d.$$

This shows that if β overlaps positively with all the α_i's then it satisfies the definition of a positive chart for the form ω which implies that ω induces $\{\alpha_i\}$. That ω as defined is a volume form is quite simple to show. Let $p \in M$ and choose β to be a positive patch around p. Then

$$\beta^*\omega = \underbrace{u}_{>0} dx^1 \wedge \cdots \wedge dx^d$$

as above. □

Proposition 7.12 (Relation between Volume Forms which Induce the same Orientation): Suppose two volume forms $\omega_1, \omega_2 \in \Omega^d(M)$ induce the same orientation on M. then there exists a strictly positive function $f \in C^\infty(M)$ such that $f(p) > 0$ for all points $p \in M$ such that

$$\omega_1 = f\omega_2$$

Proof: Since volume forms are *nowhere-vanishing* we know that f must be uniquely determined by the equation $\omega_1 = f\omega_2$. Let α be a positive chart in the common induced orientation of ω_1 and ω_2. Then

$$\alpha^*\omega_1 = \alpha^* f \cdot \alpha^*\omega_2 \longleftrightarrow u_1 dx^1 \wedge \cdots \wedge dx^d = \alpha^* f \cdot u_2 dx^1 \wedge \cdots \wedge dx^d$$

where $u_1, u_2 > 0$. Therefore

$$u_1 = \alpha^* f u_2 \implies \alpha^* f > 0 \implies f > 0$$

□

From the results of Propositions 7.10 - 7.12 follows immediate theorem whose proof is simply given by the results of the propositions

Theorem 7.13: There exists a one-to-one correspondence between volume forms on a manifold and orientations on a manifold.

Theorem 7.14 (Boundary of Orientable Manifold is Orientable): Let M be an oriented d-manifold with boundary $\partial M \neq \varnothing$. Then ∂M is orientable.

Proof: We start by defining a set of coordinate patches specifically for ∂M as restrictions of the standard set of positive patches covering M (a procedure we used in Theorem 3.5)

$$\mathrm{A} = \{\alpha|_{U\cap\partial\mathbb{H}^d} : U \cap \partial\mathbb{H}^d \to V \cap \partial M \mid \alpha \text{ a positive patch}\}.$$

Then A covers ∂M since the unrestricted set of positive patches cover M. It remains to show that any two elements of A overlap positively. Let $\alpha_i : U_i \to V_i$ with $i = 1, 2$ be elements of A such that $(V_1 \cap \partial M) \cap (V_2 \cap \partial M) \neq \varnothing$. WLOG we can set $V_1 = V_2 \equiv V$. Let the transition function of the unrestricted patches be $\Phi = \alpha_1^{-1} \circ \alpha_2 : U_2 \to U_1$. Then we can obtain the restricted version as

$$\Phi|_{U_2\cap\partial\mathbb{H}^d} = \alpha_1^{-1}|_{V_1\cap\partial M} \circ \alpha_2|_{U_2\cap\partial\mathbb{H}^d} : (U_2 \cap \partial\mathbb{H}^d) \to (U_1 \cap \partial\mathbb{H}^d)$$

We know that Φ is orientation preserving since the unrestricted patches are positive by assumption. We also have that $(U_2 \cap \partial\mathbb{H}^d) = \Phi^{-1}(U_1 \cap \partial\mathbb{H}^d)$. The consequences can be easier to understand by observing that since the d'th component of $\Phi = (\Phi_1, \ldots, \Phi_d)$ is the only component which changes under the restriction, we have

$$\Phi_d(x^1, \ldots, x^{d-1}, x^d) \geq 0 \;\text{ for } x^d \geq 0 \implies x \in \mathbb{H}^d,$$

$$\Phi_d(x^1, \ldots, x^{d-1}, 0) = 0 \;\text{ for } x^d = 0 \implies x \in \partial\mathbb{H}^d.$$

Then since points $(x^1, \ldots, x^{d-1}, 0)$ have the same form as points in $(U_2 \cap \partial\mathbb{H}^d)$, we can see that for $x \in (U_2 \cap \partial\mathbb{H}^d)$:

$$D\Phi_d(x) = \left(0, \ldots, 0, \underbrace{\frac{\partial\Phi_d}{\partial x^d}}_{\geq 0}\right)$$

$$\implies D\Phi(x) = \begin{pmatrix} D(\Phi|_{U_2 \cap \partial \mathbb{H}^d})(x) & ? \\ 0 & \frac{\partial \Phi_d}{\partial x^d} \geq 0 \end{pmatrix}$$

where we reach the fact that

$$\det[D(\Phi|_{U_2 \cap \partial \mathbb{H}^d})(x)] \cdot \frac{\partial \Phi_d}{\partial x^d} = \det[D\Phi] \implies \det[D(\Phi|_{U_2 \cap \partial \mathbb{H}^d})] > 0$$

since the right hand side is zero by assumption. Thus ∂M is orientable with orientations given by restricting an orientation on the manifold M. $\square$

Definition (Induced Orientation on the Boundary of a Manifold): In the above proof we constructed an orientation on ∂M by restricting an orientation on M. We refer to the **restricted orientation** on ∂M as the orientation obtained above and the **induced orientation** as $(-1)^d$ times the restricted one.

Definition (Alternate Definition of Orientation Preserving Diffeomorphism): Let $f : M \to N$ be a diffeomorphism between oriented manifolds with boundary $(M, \{\alpha_i\})$ and $(N, \{\beta_i\})$. We say f is orientation preserving if $\{f \circ \alpha_i\}$ are positive charts with respect to the orientation $\{\beta_i\}$.

There is an analogous result to Proposition 7.12 for two manifolds related to each other by a diffeomorphism offered as a homework problem but I feel that it is worth including the proof here as they reinforce the useful techniques we are developing for dealing with orientations and volume forms.

Proposition 7.15 (Condition for a Diffeomorphism between Oriented Manifolds to be Orientation Preserving): Let $F : M \to N$ be a diffeomorphism between two oriented d-manifolds with boundary M and N. Suppose the orientations are induced by volume forms $\omega_M \in \Omega^d(M)$ and $\omega_N \in \Omega^d(N)$ respectively. Then F is orientation-preserving if and only if there exists a positive function $u \in C^\infty(M)$ such that

$$F^* \omega_N = u \omega_M$$

Proof: Let $\xi : U_1 \subset \mathbb{H}^d \to M$ be a function that overlaps positively with the orientations of M and $\chi : U_2 \subset \mathbb{H}^d \to N$. be a function that overlaps positively with the orientations of N. Then from our study of differential forms, we know we can express the forms $\xi^* \omega_M$ and $\chi^* \omega_N$ uniquely in the form

$$\xi^* \omega_M = f dx^1 \wedge dx^2 \wedge \cdots \wedge dx^d \qquad \text{and} \qquad \chi^* \omega_N = g dx^1 \wedge dx^2 \wedge \cdots \wedge dx^d$$

where f and g must be positive functions given that ξ and χ are positive with respect to each orientation. We can equate now the wedge products of elementary forms in both expressions which gives

$$\frac{\xi^* \omega_M}{f} = \frac{\chi^* \omega_N}{g} \implies \omega_N = \left(\frac{g}{f}\right)(\chi^{-1})^* \xi^* (\omega_M) = (\chi^{-1})^* \xi^* \left(\frac{g}{f} \omega_M\right)$$

$$= (\xi \circ \chi^{-1})^* \left(\frac{g}{f} \omega_M \right).$$

Now letting $F : M \to N$ be a diffeomorphism between the manifolds with boundary we show

$$F^* \omega_N = F^* (\xi \circ \chi^{-1})^* \left(\frac{g}{f} \omega_M \right) = (\xi \circ \chi^{-1} \circ F)^* \left(\frac{g}{f} \omega_M \right)$$

for some point $p \in M$. Now

$$(F^* \omega_N)(p) = (DF(p))^* (\omega_N)_{F(p)} = \det DF(p) (\omega_N)_{(F(p))} = \det D[\xi \circ \chi^{-1} \circ F(p)] (\frac{g}{f}) (\omega_M)_{(\xi \circ \chi^{-1} \circ F(p))}.$$

Now since f, g are both positive functions by definition, we see that if we define $u(p)$ to be the function

$$u(p) = \det D[\xi \circ \chi^{-1} \circ F(p)] \frac{g}{f}$$

then we have that $F^* \omega_N = u \omega_M$ and in this form we have shown that F is orientation preserving iff u is positive ($u(p) > 0$ for all $p \in M$). □

Now we can finally define integration of forms the way we wanted to based on Proposition 7.3 without having any holes in our reasoning.

Definition (Integration of Forms over Compact Manifolds with Boundary): Let M be a compact d-manifold with boundary. Let M be covered by positive patches $\{\alpha_i : U_i \to V_i\}$ and let $\{\varphi_i\}$ be a partition of unity subordinate to $\{V_i\}$. We define the operation $\int_M : \Omega^d(M) \to \mathbb{R}$ by

$$\int_M \omega = \sum_{i=1}^{l} \int_{U_i} \alpha_i^* (\varphi_i \omega)$$

where

$$\int_{U_i} \alpha_i^* (\varphi_i \omega) = \int_{\mathbb{R}^d} u dx^1 \cdots dx^d.$$

The same definition holds for non-compact manifolds but only for $\omega \in \Omega_c^d(M)$. Note that M being compact means that $\Omega_c^\circ(M) = \Omega^\circ(M)$.

Proposition 7.16 (This Definition of Integration is Well-Defined): The definition of integration above is well-defined , in other words, it is independent of the choices of α_i's and φ_i's.

Proof: Let $\{\alpha_i, \varphi_i\}$ and $\{\beta_i, \phi_i\}$ be two choices of positive patches and partitions of unity. The definition is not affected by the addition of some α_k if we choose the corresponding φ_k to be 0. So we can assume that $\alpha_i = \beta_i$ and reduce the problem to showing how the definition is choice-independent of the partition of unity. We can show this by computation:

$$\sum_i \int_{U_i} \alpha_i^* (\varphi_i \omega) = \sum_i \int_{U_i} \alpha_i^* \left(\sum_j \phi_j \varphi_i \omega \right) = \sum_{i,j} \int_{U_i} \alpha_i^* (\phi_j \varphi_i \omega)$$

we have that the integrand is only non-zero at points in the overlap $(V_i \cap V_j)$. And

$$\alpha_i^*(\phi_j \varphi_i \omega) = (\alpha_j^{-1} \circ \alpha_i)^*(\alpha_j^*(\phi_j \varphi_i \omega))$$

thus

$$\sum_i \int_{U_i} \alpha_i^*(\varphi_i \omega) = \sum_{i,j} \int_{U_i} (\alpha_j^{-1} \circ \alpha_i)^*(\alpha_j^*(\phi_j \varphi_i \omega)) = \sum_{i,j} \int_{U_j} \alpha_j^*(\phi_j \varphi_i \omega)$$

where in the last step we have used the change of variables theorem. This is now equal to

$$\sum_j \int_{U_j} \alpha_j^*(\phi_j \underbrace{\sum_i \varphi_i}_{=1} \omega) = \sum_j \int_{U_j} \alpha_j^*(\phi_j \omega)$$

Although all we really did in this proof is insert the number 1 into the integrand and take a different number 1 out it still counts as a proof! □

Theorem 7.17 (Properties of Integration of Forms): Let M be an oriented manifold with boundary with $\omega, \eta \in \Omega^d(M)$ and $\lambda \in \mathbb{R}$, the following properties hold:

1. The integral is linear, i.e.

$$\int_M (\lambda \omega + \eta) = \lambda \int_M \omega + \int_M \eta$$

2. If $-M$ denotes the manifold with the opposite orientation then

$$\int_{-M} \omega = -\int_M \omega$$

Proof: The proof of the first property simply follows from that fact that all operations involved in the integral are linear. To prove (2) suppose that $\{\alpha_i\}$ are positive charts on M, then $\{\alpha_i \circ \tau\}$ are positive charts on $-M$ where τ is the standard orientation-reversing diffeomorphism. Then we can compute

$$\int_{-M} \omega = \sum_i \int_{\tau(U_i)} (\alpha_i \circ \tau)^* \varphi_i \omega = \sum_i \int_{\tau(U_i)} \tau^*(\alpha_i^*(\varphi_i \omega))$$
$$= \sum_i \int_{\tau(U_i)} u(-x^1, x^2, \dots, x^d)(\det D\tau) dx^1 \cdots dx^d$$

but since $\det D\tau = -1$ we have that

$$\int_{-M} \omega = -\sum_i \int_{U_i} u(x^1, \dots, x^d) dx^1 \cdots dx^d = \int_M \omega$$

□

Theorem 7.18 (Integrating a Form under the Induced Map of an Orientation Preserving Diffeomorphism): Let $F : M \to N$ be an orientation-preserving diffeomorphism between manifolds M and N. Then

$$\int_M F^* \omega = \int_N \omega$$

for all $\omega \in \Omega_c^d(N)$.

Proof: Let $\{\alpha_i\}$ be a collection of positive charts covering M. Let $\{\varphi_i\}$ be a partition of unity subordinate to $\{V_i\}$. Then $\{(F^{-1})^*\varphi_i\}$ and $\{F \circ \alpha_i\}$ serve the same purpose on the manifold N. Then we can show that

$$\int_N \omega = \sum_i \int_{U_i} (F \circ \alpha_i)^*((F^{-1})^* \varphi_i \omega) = \sum_i \int_{U_i} \alpha_i^*(F^{-1} \circ F)^* \varphi_i \cdot F^* \omega)$$

$$= \sum_i \int_{U_i} \alpha_i^*((\varphi_i F^* \omega) = \int_M F^* \omega.$$

□

Lemma 7.19 (Integrating the Differential of a Form) Let $b : \mathbb{R}^{d-1} \to \mathbb{H}^d$ be the **inclusion of the boundary** defined by the equation

$$b(x^1, \dots, x^{d-1}) = (x^1, \dots, x^{d-1}, 0).$$

Then for any $\eta \in \Omega_c^{d-1}(\mathbb{H}^d)$ we have

$$\int_{\mathbb{H}^d} d\eta = (-1)^d \int_{\mathbb{R}^{d-1}} b^* \eta.$$

Proof: Writing η in terms of elementary 1-forms:

$$\eta = \sum_i f_i dx^1 \wedge \cdots \wedge dx^{i-1} \wedge \hat{dx^i} \wedge dx^{i+1} \wedge \cdots \wedge dx^d \text{ where } f_i \in C_c^\infty(\mathbb{H}^d).$$

Let's denote the large wedge product here excluding the ith form as simply dx^{I_i}. Both sides of the equation in the lemma are linear in η so we can assume that $\eta = f dx^{I_i}$. Then the differential of η is

$$d\eta = df \wedge dx^{I_i} = \frac{\partial f}{\partial x^i}(-1)^{i-1} dx^1 \wedge \cdots \wedge dx^d.$$

Since $\eta \in \Omega_c^{d-1}(\mathbb{H}^d)$ we know that f has compact support thus we can assume

Support $f \subseteq [-L, L]^{d-1} \times [0, L]$ and $f(x^1, \dots, x^d) = 0$ whenever $x^i = \pm L$ or $x^d = L$.

Basically we are just saying that f vanishes at some arbitrary boundary on the upper-half space that we made up for convenience. Now we can go ahead and compute

$$\int_{\mathbb{H}^d} d\eta = \int_{[-L,L]^{d-1} \times [0,L]} \frac{\partial f}{\partial x^i} (-1)^{i-1} dx^1 \cdots dx^d$$

$$= \begin{cases} \int_{[-L,L]^{d-2} \times [0,L]} (-1)^{i-1} \left(\int_{[-L,L]} \frac{\partial f}{\partial x^i} dx^i \right) dx^1 \cdots \hat{dx^i} \cdots dx^d & \text{if } i \neq d, \\ \int_{[-L,L]^{d-1}} (-1)^{d-1} \left(\int_{[0,L]} \frac{\partial f}{\partial x^d} dx^d \right) dx^1 \cdots dx^{d-1} & \text{if } i = d. \end{cases}$$

$$= \begin{cases} 0 \text{ (by assumption)} & \text{if } i \neq d, \\ \int_{[-L,L]^{d-1}} (-1)^d f(x^1, \ldots, x^{d-1}, 0) dx^1 \cdots dx^{d-1} & \text{if } i = d. \end{cases}$$

Now to compute the other side of the equation we see that since b sends the x^d coordinate to 0 and fixes all others, $b^*\eta$ can only be non-vanishing if the dx^d form is the one that is excluded i.e. $i = d$

$$b^* dx^i = \begin{cases} dx^i & \text{if } i \neq d, \\ 0 & \text{if } i = d. \end{cases} \implies b^*\eta = \begin{cases} 0 & \text{if } i \neq d, \\ f(x^1, \ldots, x^{d-1}, 0) dx^1 \cdots dx^{d-1} & \text{if } i = d. \end{cases}$$

and finally

$$\int_{\mathbb{R}^{d-1}} b^*\eta = \begin{cases} 0 & \text{if } i \neq d, \\ \int_{\mathbb{R}^{d-1}} f(x^1, \ldots, x^{d-1}, 0) dx^1 \cdots dx^{d-1} & \text{if } i = d. \end{cases}$$

Sending L back out to infinity makes both sides equal. □

(Abuse of) Notation: Let $b : \partial M \to M$ be the inclusion of the boundary as defined above. Suppose M is oriented, then for $\omega \in \Omega^{d-1}(M)$ we write

$$\int_{\partial M} \omega \equiv \int_{\partial M} b^*\omega$$

Corollary to Lemma 7.19: If $\eta \in \Omega_c^{d-1}(\mathbb{R}^d)$ and we instead integrate $d\eta$ over the entire space $\mathbb{R}^d$ then the integral vanishes

$$\int_{\mathbb{R}^d} d\eta = 0 \qquad \eta \in \Omega_c^{d-1}(\mathbb{R}^d).$$

Theorem 7.20 (Stokes): Let M be an oriented manifold with boundary. Then

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

where $\omega \in \Omega_c^{d-1}(M)$ or M itself is compact.

Remark: If M has no boundary, then the analogue of Stokes' Theorem becomes

$$\int_M d\omega = 0 \qquad \omega \in \Omega_c^{d-1}(M)$$

Proof: Let $\{\alpha_i\}$ be a collection of positive charts covering M. Choose $\{\varphi_i\}$ be a partition of unity subordinate to the covering of M, $\{V_i\}$ such that $\omega \in \Omega_c^{d-1}(M)$can be written in the form

$$\omega = \sum_i \varphi_i \omega.$$

Now since Support ω is compact, only finitely many terms $\varphi_i\omega$ in the sum are non-zero. Since the d operator and the integrals are linear it suffices to show that the theorem holds in the special case where (Support ω) $\cap M$ can be covered by open sets where finitely many partition functions φ_i are non-zero. Now let $\alpha : U \to V$ be the positive chart such that V is the set covering (Support ω) $\cap M$. Then we can compute

$$\int_M d\omega = \int_{\mathbb{H}^d} \alpha^* d\omega = \int_{\mathbb{H}^d} d(\alpha^*\omega) \overset{\text{by Lemma 7.19}}{=} (-1)^d \int_{\mathbb{R}^{d-1}} ((b|_{\mathbb{H}^d})^* \alpha^* \omega)$$

where $b|_{\mathbb{H}^d}$ is the inclusion of the boundary $b : \mathbb{R}^{d-1} \to \mathbb{H}^d$. We can use the property of the induced transformation to rewrite the composition in the integrand as

$$(-1)^d \int_{\mathbb{R}^{d-1}} b^*(\alpha^*(\omega)) = (-1)^d \int_{\mathbb{R}^{d-1}} (\alpha \circ b)^* \omega.$$

But $\alpha \circ b : b^{-1}(U) \to V$ is the restricted chart corresponding to α on ∂M, and $b^*\omega$ has support covered by V. In particular, we now have that there exists a chart $\overline{\alpha} : b^{-1}(U) \to \partial M$ such that $b|_M \circ \overline{\alpha} = \alpha \circ b$.

Let ∂M_{res} be the boundary ∂M with the restricted orientation, so that $\alpha \circ b$ is a positive chart. Then we can compute the right hand side as

$$\int_{\partial M} (b|_M)^* \omega = (-1)^d \int_{\partial M_{\text{res}}} (b|_M)^* \omega = (-1)^d \int_{\mathbb{R}^{d-1}} \overline{\alpha}^* (b|_M^* \omega)$$

$$= (-1)^d \int_{\mathbb{R}^{d-1}} (b|_M \circ \overline{\alpha})^*(\omega) = (-1)^d \int_{\mathbb{R}^{d-1}} (\alpha \circ b)^* \omega = (-1)^d \int_{\mathbb{R}^{d-1}} b^*(\alpha^*(\omega)).$$

□

Remark (For 1-Manifolds): If M is a 1-manifold with boundary, then an orientation on ∂M is the choice of a function $\sigma : \partial M \to \{\pm 1\}$ as previously discussed. We shall define this choice as follows:

If there exists a positive patch $\alpha : [0, \varepsilon] \subset \mathbb{H}^1 \to M$, then let $\sigma(\alpha(0)) = -1$, similarly if there exists a positive patch from the negative reals $\alpha : [-\varepsilon, 0] \subset \mathbb{R}^{\leq 0} \to M$ then we define $\sigma(\alpha(0)) = 1$.

Now if we define then integration of the 0-form f over ∂M as

$$\int_{\partial M} f = \sum_{p \in \partial M} \sigma(p) f(p).$$

then Stokes' theorem holds.

For example $M = [0, 1]$ means that

$$\int_0^1 \frac{\partial f}{\partial x} dx = \int_M df = \int_{\partial M} f = f(1) - f(0).$$