Introduction to Manifolds & Analysis

Ian Poon

August 2024

Selected theorems from [1]

Contents

1 Differentiation in n dimensions

Definition 1

Recall in 1D

$$f'(x) = \lim_{t\to 0} \frac{f(x+t) - f(x)}{t} = \lambda$$

equivalently we write

$$f'(x) = \lim_{t\to 0} \frac{f(x+t) - f(x) - \lambda t}{t} = 0$$

Definition 2

Give an open set U in $\mathbb{R}^m$ a map $f : U \to \mathbb{R}^n$, a point $a \in U$ and a point $u \in \mathbb{R}^n$ the **directional derivative** of f in the direction of u at a is

$$Df(\mathbf{a}) \cdot \mathbf{u} = D_u f(\mathbf{a}) = \lim_{t\to 0} \frac{f(\mathbf{a} + t\mathbf{u}) - f(\mathbf{a})}{t}$$

where the $\cdot$ here refers to matrix product and t is some scalar.In particular, if we calculate directional derivatives with respect to the direction of standard basis vectors e_i where $i = 1, \dots, n$ of $\mathbb{R}^n$ we denote it as

$$Df(\mathbf{a}) \cdot \mathbf{e}_i = D_i f(\mathbf{a}) = D_{\mathbf{e}_i} f(\mathbf{a}) = \frac{\partial f(\mathbf{a})}{\partial \mathbf{x}_i} = \lim_{t\to 0} \frac{f(\mathbf{a} + t\mathbf{e}_j) - f(\mathbf{a})}{t}$$

The name directional comes from the fact that is a matrix product of some $Df(\mathbf{a})$ and a column vector that specifies our "direction" as seen above. Specifically the hypervector $Df(\mathbf{a}) = (D_1 f(\mathbf{a}), \dots, D_m f(\mathbf{a}))$ will be discussed more below

Example 3

You can see from this example the motivation behind the defintion above

$$Df(a) \cdot e_2 = (D_1 f(a), D_2 f(a), D_3 f(a)) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = D_2 f(a)$$

where the $\cdot$ here refers to matrix product. We will continue to use this convention for the rest of the notes

Definition 4

Give an open subset U of $\mathbb{R}^m$ a map $f : U \to \mathbb{R}^n$ and a point $a \in U$ the function f is **differentiable** at a if there exists a $n \times m$ matrix B of linear transformation from $\mathbb{R}^m \to \mathbb{R}^n$ such that for every $h \in \mathbb{R}^m - \{0\}$

$$\frac{f(a+h) - f(a) - Bh}{|h|} \to 0$$

as $h \to 0$. We will denote **derivative** of f at a by

$$Df(a) = B$$

and that B is a unique matrix. We will how this definition is motivated below

Theorem 5

If f is differentiable at a then all directional derivatives of f at a exist

$$f'(\mathbf{a}; \mathbf{u}) = Df(\mathbf{a}) \cdot \mathbf{u}$$

Proof. Let $B = Df(a)$.Let t be a scalar. Suppose for every $\mathbf{u}$ the function f is differentiable at a so

$$\begin{aligned} \frac{f(a+t\mathbf{u}) - f(a) - B \cdot t\mathbf{u}}{|t\mathbf{u}|} &= \frac{t}{|t\mathbf{u}|} \frac{f(a+t\mathbf{u}) - f(a) - B \cdot t\mathbf{u}}{t} \\ &= \frac{1}{|t|} \frac{t}{|\mathbf{u}|} \left(\frac{f(a+t\mathbf{u}) - f(a)}{t} - B \cdot \mathbf{u} \right) \\ &\to 0 \end{aligned}$$

as $t \to 0$, regardless of what $\mathbf{u}$ is so clearly

$$\lim_{t \to 0} \frac{f(a+t\mathbf{u}) - f(a)}{t} = B \cdot \mathbf{u}$$

exists for all u as desired

Theorem 6

If f is differentiable at a then f is continuous at a

Proof. Similar to the what we have done previously in `Jeri Lebl Introduction Real analysis` we do

$$\lim_{h \to 0} f(a+h) - f(a) = \left(\lim_{t \to 0} \frac{f(a+h) - f(a) - Bh}{|h|} \right) \left(\lim_{h \to 0} |h| \right) + Bh$$

so $\lim_{h \to 0} f(a+h) = f(a)$ which shows continuity

Fact 7 (Jacobian Matrix)

Let $Df(a)$ be a matrix of a linear map $\mathbb{R}^m \to \mathbb{R}^n$ and f be the map $f : U \to \mathbb{R}^n$ where $U \subseteq \mathbb{R}^m$. We can write every output of $f(x)$ in our codomain $\mathbb{R}^n$ as

$$f(x) = f_1(x) + \ldots + f_n(x) = f(x)e_1 + \ldots + f(x)e_n$$

where each component function is map $f_i : U \to \mathbb{R}$. Therefore our $n \times m$ transformation matrix $Df(x)$ which we call the **Jacobian Matrix** is defined by

$$Df(a) = [D_j f_i(a)]$$

where columns $j = 1, \ldots, m$ and rows $i = 1, \ldots, n$

Example 8 (Linear functions vs Jacobian)

Now, suppose $\mathbf{f}$ is a linear function, i.e., $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$ where A is a constant $m \times n$ matrix. Then for any $\mathbf{x}$, we have:

$$\mathbf{f}(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

This results in:

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \sum_{j=1}^n a_{1,j}x_j \\ \sum_{j=1}^n a_{2,j}x_j \\ \vdots \\ \sum_{j=1}^n a_{m,j}x_j \end{bmatrix}$$

The i-th component of $\mathbf{f}(\mathbf{x})$ is:

$$f_i(\mathbf{x}) = \sum_{j=1}^{n} a_{i,j}x_j$$

To find the partial derivative of f_i with respect to x_j, we calculate:

$$\frac{\partial f_i(\mathbf{x})}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\sum_{k=1}^{n} a_{i,k}x_k\right)$$

Since $a_{i,k}$ are constants (because A is a constant matrix), this simplifies to:

$$\frac{\partial f_i(\mathbf{x})}{\partial x_j} = a_{i,j}$$

Theorem 9

Suppose that all of the partial derivatives $\frac{\partial f}{\partial x_j}$ in the *jacobian matrix* exist at all points $x \in U$ and that all of the partial derivatives are continuous at $x = a$ then f is differentiable at a

Proof. Consider

$$\mathbf{p}_0 = \mathbf{a} = (\mathbf{a}e_1 + \mathbf{a}e_2 + \dots \mathbf{a}e_n) \tag{1}$$
$$\mathbf{p}_1 = (\mathbf{a}e_1 + \mathbf{a}e_2 \dots \mathbf{a}e_n) + \mathbf{h}e_1 \tag{2}$$
$$\mathbf{p}_2 = (\mathbf{a}e_1 + \mathbf{a}e_2 \dots \mathbf{a}e_n) + \mathbf{h}e_1 + \mathbf{h}e_2 \tag{3}$$
$$\vdots \tag{4}$$
$$\mathbf{p}_n = \mathbf{a} + \mathbf{h} = \mathbf{a} + (\mathbf{h}e_1 + \dots + \mathbf{h}e_n) \tag{5}$$

Then we know that

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = f(\mathbf{p}_n) - f(\mathbf{p}_0) = \sum_{j=1}^{n} f(\mathbf{p}_j) - f(\mathbf{p}_{j-1})$$

by mean value theorem for each term in this sum we can find $0 \leq \xi_j \leq \mathbf{h}e_j$ such that

$$f(\mathbf{p}_j) - f(\mathbf{p}_{j-1}) = Df(\mathbf{p}_{j-1} + \xi_j)\mathbf{h}e_j$$

therefore we can write

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - Df(\mathbf{a})h = \sum_{j=1}^{n} Df(\mathbf{p}_{j-1} + \xi_j)\mathbf{h}e_j - Df(\mathbf{a})\mathbf{h}e_j \tag{6}$$
$$\frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - Df(\mathbf{a})h}{|h|} = \frac{1}{|h|}\sum_{j=1}^{n} Df(\mathbf{p}_{j-1} + \xi_j)\mathbf{h}e_j - Df(\mathbf{a})\mathbf{h}e_j \tag{7}$$

So $h \to 0$ so will all $\mathbf{h}e_j$ as well as ξ_j too. Moreover all the parts of $\mathbf{h}$ in $\mathbf{p}_j$ will disappear too so every $\mathbf{p}_j \to \mathbf{a}$. Therefore all the partial derivatives are continuous since $Df(\mathbf{p}_{j-1} + \xi_j)\mathbf{h}e_j \to Df(\mathbf{a})\mathbf{h}e_j$ and the RHS goes to zero as $\mathbf{h} \to 0$ as desired

Definition 10

$\mathcal{C}^r$ means continuous up to the r order, for example both $f(x)$ and $f'(x)$ are continuous for $\mathcal{C}^1$. More precisely Given $U \subseteq \mathbb{R}^n$ and $f : U \to \mathbb{R}$ we define

$$f \in \mathcal{C}^1(U) \Leftrightarrow \frac{\partial f}{\partial x_i}, i, \dots, n \text{ exist and continuous on all points } x \in U$$

$$\vdots$$

$$f \in \mathcal{C}^k(U) \Leftrightarrow \frac{\partial f}{\partial x_i} \in \mathcal{C}^{k-1}, i, \dots, n \text{ exist and continuous on all points } x \in U$$

$$f \in \mathcal{C}^\infty(U) \Leftrightarrow \frac{\partial f}{\partial x_i}, i, \dots, n \text{ of all orders exist and continuous on all points } x \in U$$

Theorem 11

Partial derivatives can be taken in any order that is

$$\frac{\partial}{\partial x_2}\frac{\partial}{\partial x_1} f(a_1, a_2) = \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2} f(a_1, a_2)$$

The idea you want to define some difference function that allows you to find the sum of all possible differences in a function so as to "approximate" $f(a+h) - f(a)$ when h is small. This means I want

1. $f(a_1 + h1, s)$ vs $f(a_1, s)$ for every possible permutation s
2. $f(s, a_2 + h_2)$ vs $f(s, a_2)$ for every possible permutation s

Then notice the following does have all such differences

$$\begin{aligned}
\Delta(h) &= f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) - f(a_1 + h_1, a_2) + f(a_1, a_2) && (1)\\
&= (f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2)) - (f(a_1 + h_1, a_2) - f(a_1, a_2)) && (2)\\
&= \phi(a_2 + h_2) - \phi(a_2) && (3)\\
&= \phi'(c_2)h_2, \quad (a_2 \le c_2 \le a_2 + h_2) && (4)\\
&= \frac{\partial (f(a_1 + h_1, c_2) - f(a_1, c_2))}{\partial x_2} h_2 && (5)\\
&= \frac{\partial \left(\frac{\partial f(c_1, c_2)}{\partial x_1} h_1 \right)}{\partial x_2} h_2, \quad (a_1 \le c_1 \le a_1 + h_1) && (6)\\
&= \frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} d(c_1, c_2) h_1 h_2 && (7)
\end{aligned}$$

You can do the analagously by defining $\phi(a_1 + h_1) - \phi(a_1)$. Just group the terms containing $a_1 + h_1$ and a_1 into 2 groups instead and use d instead of c because they don't necessarily have to be the same, just need to be in the interval as specified by mean value theorem. And then you will achieve

$$\Delta(h) = \frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} f(c_1, c_2) h_1 h_2 = \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f(d_1, d_2) h_2 h_1$$

but as $h \to 0$ both c_1, c_2 and d_1, d_2 go to a_1, a_2 as desired. Now by induction, any further partial differentials of $\Delta(h)$ will also follow this property, for example suppose we let from the previous case $\Delta(h \to 0) = F(a_3, a_4)$ where a_3, a_4 were existing variables we just treated as constants/didn't bother with. We can prove in the same way that

$$\frac{\partial}{\partial x_4} \frac{\partial}{\partial x_3} F(a_3, a_4) = \frac{\partial}{\partial x_3} \frac{\partial}{\partial x_4} F(a_3, a_4)$$

Theorem 12 (Lebniz Integral Rule)
Let f be continuously differentiable real functions on some region R of the (x, t) plane. Then for all $(x, y) \in \mathbb{R}$

$$\frac{d}{dt} \int_{a(t)}^{b(t)} f(x, t) dx = f(b(t), t) \frac{db}{dt} - f(a(t), t) \frac{da}{dt} + \int_{a(t)}^{b(t)} \frac{\partial f(x, t)}{\partial t} dx$$

We also call this **differentiation under the integral**. It follows that if a and b are constants the first 2 terms on the RHS vanishes.

Proof. Consider

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(x,t)\,dx = \lim_{h\to 0}\frac{1}{h}\left(\int_{a(t+h)}^{b(t+h)} f(x,t+h)\,dx - \int_{a(t)}^{b(t)} f(x,t)\,dx\right) \tag{1}$$

$$= \lim_{h\to 0}\frac{1}{h}\left(\int_{a(t)}^{b(t)} \left(f(x,t+h)-f(x,t)\right)\,dx\right. \tag{2}$$

$$\left.+\int_{b(t)}^{b(t+h)} f(x,t+h)\,dx - \int_{a(t)}^{a(t+h)} f(x,t+h)\,dx\right) \tag{3}$$

$$= \lim_{h\to 0}\frac{1}{h}\left((b(t+h)-b(t))\,f(\xi_b,t+h)\right. \tag{4}$$

$$\left.-\,(a(t+h)-a(t))\,f(\xi_a,t+h) + \int_{a(t)}^{b(t)} \left(f(x,t+h)-f(x,t)\right)\,dx\right) \tag{5}$$

$$= \lim_{h\to 0}\left(\frac{b(t+h)-b(t)}{h}\cdot f(\xi_b,t+h)\right. \tag{6}$$

$$\left.-\,\frac{a(t+h)-a(t)}{h}\cdot f(\xi_a,t+h) + \frac{1}{h}\int_{a(t)}^{b(t)} \left(f(x,t+h)-f(x,t)\right)\,dx\right). \tag{7}$$

$$= f(b(t),t)\frac{db}{dt} - f(a(t),t)\frac{da}{dt} + \int_{a(t)}^{b(t)} \frac{\partial f(x,t)}{\partial t}dx \tag{8}$$

(4) and (5) is consequence of mean value theorem for integrals of a continuous real function on closed interval that can be seen from

$$\int_b^a f(x)dx = f(k)(b-a)$$

for some $k \in [a,b]$. (7) follows when we can define

$$f_n = \frac{f(x,t+h_n)-f(x,t)}{h_n}$$

where $h_n \to 0$ as $n \to 0$ hence

$$\lim_{h\to 0} f = \lim_{n\to\infty} f_n = \frac{\partial f(x,t)}{\partial t}dx$$

and that suprenum of $f(x,t)$ over R_T which is the set of all x for a given t exists so $|f_n| < \infty$ for a given t. Thefore by dominated convergence theorem we have the conclusion in (8)

Theorem 13 (Multivar Chain Rule)
Let U and v be open sets in $\mathbb{R}^n$. Consider maps $f : U \to V$ and $g : V \to \mathbb{R}^k$. Choose $a \in U$ and let $b = f(a)$. If f is differentiable at a and g is differentiable at b then $g \circ f$ is differetiable at a and the derivative is

$$(Dg \circ f)(a) = (Dg)(b) \cdot Df(a)$$

where the $\cdot$ here refers to the *matrix product*

Proof. Consider for f at a

$$\Delta(h) = f(a+h) - f(a) \tag{1}$$

$$F(h) = \frac{f(a+h)-f(a)-Df(a)h}{|a|} \tag{2}$$

$$\Delta(h) = Df(a)h + |h|\,F(h) \tag{3}$$

We do the same for g at b

$$G(k) = \frac{g(b+k) - g(b) - Dg(b)k}{|k|} \tag{4}$$
$$g(b+k) - g(b) = Dg(b)k + |k|\, G(k) \tag{5}$$

Now consider $g \circ f$ at a

$$\begin{aligned} g \circ f(a+h) - g \circ f(a) &= g(f(a+h)) - g(f(a)) && (6) \\ &= g(b + \Delta(h)) - g(b) && (7) \end{aligned}$$

since $f(a) = b$ and $f(a+h) = f(a) + \Delta(h) = b + \Delta(h)$ Therefore using the form in (5) then subbing in $k = \Delta(h)$ from (3) we have

$$\begin{aligned} g \circ f(a+h) - g \circ f(a) &= Dg(b)\Delta(h) + |\Delta(h)|\, G(\Delta(h)) && (8) \\ &= Dg(b)(Df(a)h + |h|\, F(h)) + |\Delta(h)|\, G(\Delta(h)) && (9) \\ &= Dg(b) \circ Df(a)h + |h|\, Dg(b)F(h) + |\Delta(h)|\, G(\Delta(h)) && (10) \\ & && (11) \end{aligned}$$

Therefore we have

$$\frac{g \circ f(a+h) - g \circ f(a) - Dg(b) \circ Df(a)h}{|h|} = Dg(b)F(h) + \frac{|\Delta(h)|}{|h|} G(\Delta(h)) \tag{12}$$

As $h \to 0$, $F(h) \to 0$ since f differentiable at a and term is in front $Dg(b)$ is just a constant. $G(\Delta(g)) \to 0$ since g differentiable at b. The conclusion follows when we can prove the below

Lemma 14

The following is bounded
$$\frac{|\Delta(h)|}{|h|}$$

Proof. Consider
$$\Delta(h) = Df(a)h + |h|\, F(h)$$
then we have
$$\frac{\Delta(h)}{|h|} = Df(a)\frac{h}{|h|} + F(h)$$
It is clear the RHS goes to zero as $h \to 0$ therefore even if you take to absolute value on everything to get $|\Delta(h)|$ in the numerator, it will likewise go to zero too by definiteness of sup norm(see below).

Theorem 15

If $f : U \to V \subseteq \mathbb{R}^n$ is a $\mathcal{C}^r$ map and $g : V \to \mathbb{R}^p$ is a $\mathcal{C}^r$ map then $g \circ f : U \to \mathbb{R}^p$ is a $\mathcal{C}^r$ map

Proof. Case $r = 1$. The rest can be seen by induction. Recall `rudin`, we know that the composite function $Dg(f(x))$ will be continuous on U. Clearly $Df(x)$ is contiuous on U by definition. Consider

$$Dg \circ f(x) = Dg(f(x)) \cdot Df(x)$$

We see that

$$Dg(f(x)) \sim \left[\frac{\partial g_i}{\partial x_i}(f(x))\right]$$

and

$$Df(x) \sim \left[\frac{\partial f_i}{\partial x_i}\right]$$

so $\frac{\partial f_i}{\partial x_i}$ Therefore we have all entries in the matrix products are products of continuous functions on U. Therefore $Dg \circ f \in \mathcal{C}^1$

Theorem 16 (Mean Value Theorem)

Let A be open in $\mathbb{R}^m$ and let $f : A \to \mathbb{R}$ be differentiable on A.

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = Df(\mathbf{c}) \cdot \mathbf{h}$$

where c is a point on the line segment $a + th, 0 \leq t \leq 1$

Proof. Set $\phi(t) = f(\mathbf{a} + t\mathbf{h})$. Then ϕ is defined for t in an open interval about $[0, 1]$.Being composite of differentiable functions, by chain rule we have

$$\phi'(t) = Df(\mathbf{a} + t\mathbf{h}) \cdot \mathbf{h} = D_{\mathbf{h}}f$$

Mean value theorem in 1D implies that

$$\phi(1) - \phi(0) = \phi'(t_0) \cdot 1$$

so we may rewrite above as

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = Df(\mathbf{a} + t_0\mathbf{h}) \cdot \mathbf{h}$$

Definition 17 (Diffeomorphism)

Let U and V be open sets n $\mathbb{R}^n$ and $f : U \to V$ a $\mathcal{C}^r$ map. The map f is a **diffeomorphism** if it is bijective and $f^{-1} : V \to U$ is also $\mathcal{C}^r$

Definition 18 (Convex)

$U \subseteq \mathbb{R}^n$ is **convex** if

$$a, b \in U \quad \Rightarrow \quad (1-t)a + tb \in U$$

for all $0 \leq t \leq 1$

Definition 19

Given an open set $U \in \mathbb{R}^n$ a map $f : U \to \mathbb{R}^k$ a point $a \in U$ recall that we defined the deriviate

$$Df(a) \sim \left[\frac{\partial f_i}{\partial x_j}(a)\right]$$

We now define the norm to be

$$|Df(a)| = \sup_{i,j}\left|\frac{\partial f_i}{\partial x_j}(a)\right|$$

Definition 20

Let $x \in \mathbb{R}^n$ The **euclidean norm** of x is

$$||x|| = \sqrt{x_1^2 + \ldots x_n^2}$$

The **sup norm** of x is

$$|x| = \sup_i |x_i|$$

In fact they are related by

$$|x-y| \leq ||x-y|| \leq \sqrt{n}\,|x-y|$$

since

$$\sup_i |x_i - y_i| = \sqrt{(x_k - x_k)^2} = |x-y| \leq ||x-y|| = \sqrt{(x_1-y_1)^2 + \ldots (x_n-y_n)^2}$$

where k is the corresponding i from the sup and that

$$\sqrt{n}\,|x-y| = \sqrt{n\,|x-y|^2}$$

And that the sup norm is $||x||_1$ while euclidean norm is $||x||_2$ in precise terminology.

Definition 21

Let A be a subset of metric space X. The **interior** of A is defined by

$$\text{Int } A \equiv (\overline{A^c})^c$$

The **exterior** of A is defined by

$$\text{Ext } A \equiv \text{Int } A^c$$

The **boundary** of A is defined by the points not in the exterior or interior that is

$$\text{Bd } A \equiv X - (\text{Ext } A \cup \text{Int } A)$$

To visualize this consider that

- $A =$Bd $A +$ Int A
- $A^c =$Ext A which is an open set since no boundary points
- $\overline{A^c} =$Ext $A +$ Bd A
- $(\overline{A^c})^c =$Int A which is an open set since no boundary points

Because the boundary points of A are limit points of both A and A^c

1.1 inverse function theorem

Fact 22
Let U and V be open sets in $\mathbb{R}^n$ and let $f : U \to V$ be a $\mathcal{C}^1$ *diffeomorphism*. So let $f^{-1} = g$. Let $b = f(a)$.Then by chain rule of differentials earlier we have

$$D(\mathrm{Id}) = D(g \circ f(a)) = Dg(b) \cdot Df(a) = I$$

where as usual $\cdot$ indicates the matrix product. Therefore the total derivative is bijective. That is

$$\det Df(u) \neq 0$$

for all $u \in U$.

However the reverse claim is not necessarily true. The next theorem tells if $f : U \to V$ is a $\mathcal{C}^r$ map and $Df(a)$ is bijective for some $a \in U$, there there exists a neighbourhood where f is a **local diffeomorphism**.

Theorem 23 (Inverse function theorem)
Let U, V be an open set in $\mathbb{R}^n$, $f : U \to V$ a $\mathcal{C}^r$ map and $a \in U$. If $Df(a) : \mathbb{R}^n \to \mathbb{R}^n$ is bijective then there exists a neighbourhood U_a of a in U and neighbourhood V_b of $b \in V$ in $\mathbb{R}^n$ such that $f|U_a : U_a \to V_b$ is a $\mathcal{C}^r$ diffeomorphism

In other words if the derivative of a continuous differentiable function at a point a is bijective there exists a function that is a diffeomorphism of a neighbourhood of a at V

Lemma 24
Let U be open in $\mathbb{R}^n$ and $F : U \to \mathbb{R}^k$ be a $\mathcal{C}^1$ mapping. Also assume that U is convex. Suppose that $|Df(a)| \leq c$ for all $A \in U$. Then for all $x, y \in U$

$$|f(x) - f(y)| \leq nc\,|x - y|$$

Remark 25. *In many alternative proofs to the inverse function theorem(IFT) people often replace this with what is known as the* **contraction mapping/Banach Fixed Point theorem**. *However the lemma we used here still follows the same logic as the classic textbook approach - which involves doing some kind of bound then normalizing the determinant as we will do so below.*

Proof. Consider any $x, y \in U$. The mean value theoerem says for eery i there exists a point c on the line joining x to y such that

$$f_i(x) - f_i(y) = \sum_j^n \frac{\partial f_i}{\partial x_i}(d_i)(x_j - y_j)$$

It is mean *value* theorem that's why we must restrict f to a single dimension because "values" of $f(x)$ are actually in $\mathbb{R}^k$. Hence it follows that

$$\begin{aligned} |f_i(x) - f_i(y)| &\leq \sum_j^n c\,|x_j - y_j| \\ &\leq nc\,|x - y| \end{aligned}$$

Because this is true for any i, taking the sup on the left side wont change anything. Recall we are using sup norm not euclidean norm.

Lemma 26

Let U be open in $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ is a $\mathcal{C}^1$ map. Suppose f takes a minimum value at some point $b \in U$ then

$$\frac{\partial f}{\partial x_i}(b) = 0, i = 1, \ldots, n$$

Proof. We reduce to the one variable result. Let $b = (b_1, \ldots, b_n)$. Let $\phi(t) = f(b_2, \ldots, b_{i-1}, t, b_{i+1}, \ldots, b_n)$ which is $\mathcal{C}^1$ and has a minimum at b_i since $f(b)$ is minimum.. Then from one-variable calculus this implies that $\frac{\partial \phi}{\partial t}(b_i) = 0$. Therefore this must apply to all i or it would contradict the fact that $f(b)$ is indeed the minimum

Lemma 27

Let U be an open map in $\mathbb{R}^n$. Let $f(x) \in \mathcal{C}^1$ be map $f : U \to \mathbb{R}$. and $g(x) = x - f(x)$ be map $g : U \to \mathbb{R}^n$. Let R_δ be a ball centred around 0 with radius δ. Given

$$a = 0, \quad , f(a) = 0 \quad Df(0) = I$$

for every $\varepsilon > 0$ there exists $\delta > 0$ such that for any $x, y \in R_\delta$ we have

$$|g(x) - g(y)| < \varepsilon |x - y|$$

Proof. First $Dg(x) = I - Df(x) \Rightarrow Dg(0) = I - I = 0$. Thus $Dg(0) = 0$ by continuity to 1st order implies for every $\varepsilon > 0$ there exists $\delta > 0$ where

$$|Dg(x) - Dg(0)| = |Dg(x)| < \varepsilon \quad |x - 0| < \delta$$

Therefore we can find $x \in B_\delta$ such that $|Dg(x)| \leq \frac{\varepsilon}{n}$. Hence by 26

$$|g(x) - g(y)| < nc |x - y|$$

where $c = \frac{\varepsilon}{n}$ gives the desired result ☐

By triangle inequality and the previous lemmas

$$|x - y| \leq |g(x) - g(y)| + |f(x) - f(y)|$$
$$(1 - \varepsilon) |x - y| \leq |f(x) - f(y)|$$

Choosing $\varepsilon = \frac{1}{2}$ gives

$$|x - y| \leq 2 |f(x) - f(y)|$$

$f : R_\delta \to \mathbb{R}^n$ is now one to one ☐

At we know that at the centre of the ball $\det(Df(0)) = 1$ since $Df(0) = I$. You will learn soon that $\det(Df(x))$ is a change of variable "scaling" in this case from $R_\delta \to \mathbb{R}^n$. By the continuity of f we can then choose a δ where

$$\det(Df(x)) > \frac{1}{2}$$

Lemma 28

If $y \in B_{\frac{\delta}{4}}$ then there exists a point $c \in R_\delta$ such that $f(c) = y$.

Proof. We begin by letting $h : \overline{R}_\delta \to \mathbb{R}$ be map defined by $h(x) = ||f(x) - y||^2$. To find a case where $f(c) = y$ for some $c \in R_\delta$ for every $y \in B_{\frac{\delta}{4}}$ we must be able find x where h is a minimum and this minimum is 0 so $f(x) = y$ as desired.

The domain $\overline{R}_\delta$ is compact so h has a minimum at some point $c \in \overline{R}_\delta$. We need to ensure it is definitely in $\{R\}_\delta$ as it may or may not be in the boundary. Hence observe what happens at the boundary $x \in \overline{R}_\delta, |x| = \delta$. This implies

$$||f(x) - f(0)|| \geq \frac{\delta}{2}$$

by the previous lemma. Now relating it to h we have

$$||f(x) - y|| \geq \frac{\delta}{4}$$

since $y \in R_{\frac{\delta}{4}}$. We need to prove that this is not the minimum, that is we have to find an interior point where $h(x) < \left(\frac{\delta}{4}\right)^2$. A natural guess is try 0

$$||f(0) - y|| = ||y|| < \frac{\delta}{4}$$

by definition of being in ball where $y \in R_{\frac{\delta}{4}}$. Hence clearly the minimum must be in the interior. We now need to prove that the minimum of h is indeed zero. Since c is the minimum point we have

$$\frac{\partial h}{\partial x_j}(c) = 2\sum_{i-1}^{n}(f_i(c) - y_i)\frac{\partial f_i}{\partial x_j}(c) = 0$$

where $h(x) = \sum_{i-1}^{n}(f_i(c) - y_i)^2$ by definition of euclidean norm. However from our above lemme, we have assumed that

$$\det(Df(x)) > 0, x \in R_\delta$$

That is the diagonals this matrix cannot be zero. Therefore for every fixed j when $i = j$

$$\frac{\partial h}{\partial x_j}(c) = 2\sum_{j-1}^{n}(f_j(c) - y_j)\frac{\partial f_j}{\partial x_j}(c) = 0 \quad \Rightarrow \quad (f_j(c) - y_j) = 0$$

Since we are able to loop through all j we can prove that every

$$(f_j(c) - y_j) = 0$$

so $f(c) = y$ as desired. $\square$

To continue with our proof for inverse function theorem, we need to prove the final two properties, that is $f^{-1} : V \to U$ is continuous and f^{-1} is differentiable at 0

Lemma 29

$f^{-1} : V \to U$ is continuous

Proof. Let $x = f^{-1}(a)$ and $y = b^{-1}$. Then from the above lemmas we have

$$|a - b| = \left| f(x) - f(y) \geq \frac{x - y}{2} \right|$$

so

$$|a - b| \geq \frac{1}{2}\left|f^{-1}(a) - b^{-1}\right|$$

so clearly when $\left|f^{-1}(a) - b^{-1}\right| < \varepsilon$ we can let $|a - b| < \delta = \varepsilon$ so continuity is proven.

Lemma 30
f^{-1} is differentiable at 0 and $Df^{-1}(0) = I$

Proof. We need to show that
$$\frac{f^{-1}(0+k) - f^{-1}(0) - Df^{-1}(0)k}{|k|} \to 0, k \to 0$$
where $Df^{-1}(0) = I$. Make the substitution $h = f^{-1}(k)$ and use $h \to 0$ instead so as use the fact that f is differentiable at 0 to conclude the proposition. □

Finally, the last part of the inverse function theorem.

Corollary 31
Given that we have proven the case of inverse function for $f|U_0 : U_0 \to V_0$ where $f(0) = 0, 0 \in U$ and $f \in \mathcal{C}^1$,generalize it $f|U_a : U_a \to V_b$ and $f \in \mathcal{C}^r$

Proof. Let the map $f : U \to V$ map $a \in U$ to $b \in V$ where $U, V \in \mathbb{R}^n$. That is, $f(a) = b$. Define
$$U' = \{x - a : x \in U\}$$
Define $f' : U' \to \mathbb{R}^n$ by $f_1(x) = f(x + a) - b$ so that $f_1(0) = 0$ and $Df_1(0) = Df(a)$(use chain rule the definition of f_1 to see this). Therefore we knkow $Df_1(a)$ is bijective as well since by assumption $Df(0)$ is bijective. By the previous lemmas we know there exists
$$f|U_0' : U_0' \to V_0$$
where it is a diffeomorphism on U'. Now replace the variables back to its original form. Consider that
$$f_1(0) = 0 \Rightarrow f(a) - b = 0$$
and
$$0 \in U' \Rightarrow a \in U$$
To generalize it to $\mathcal{C}^r$ just do by induction...

1.2 Implicit function theorem

we now show that we may derive the implicit function theorem using the inverse function theorem.

Theorem 32 (Implicit Function Theorem)
Let $F : \mathbb{R}^{n+m} \to \mathbb{R}^m$ be a continuously differentiable function ($\mathcal{C}^1$) such that:

1. $F(x_0, y_0) = 0$, where $x_0 \in \mathbb{R}^n$ and $y_0 \in \mathbb{R}^m$,
2. The Jacobian matrix $\frac{\partial F}{\partial y}(x_0, y_0)$ is invertible.

Then there exist:

- Open neighborhoods $U \subset \mathbb{R}^n$ around x_0,
- Open neighborhoods $V \subset \mathbb{R}^m$ around y_0,

and a unique continuously differentiable function $g : U \to V$ such that:

$$F(x, g(x)) = 0 \quad \text{for all } x \in U.$$

Proof. Define $G : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ by:

$$G(x, y) = \begin{bmatrix} x \\ F(x, y) \end{bmatrix}.$$

We compute the Jacobian matrix $DG(x, y)$ of G at (x_0, y_0):

$$DG(x_0, y_0) = \begin{bmatrix} I_n & 0 \\ \frac{\partial F}{\partial x}(x_0, y_0) & \frac{\partial F}{\partial y}(x_0, y_0) \end{bmatrix},$$

where:

- I_n is the $n \times n$ identity matrix,
- $\frac{\partial F}{\partial x}(x_0, y_0)$ is the $m \times n$ matrix of partial derivatives of F with respect to x,
- $\frac{\partial F}{\partial y}(x_0, y_0)$ is the $m \times m$ matrix of partial derivatives of F with respect to y.

Since $DG(x_0, y_0)$ is block lower triangular, its determinant is:

$$\det(DG(x_0, y_0)) = \det(I_n) \cdot \det\left(\frac{\partial F}{\partial y}(x_0, y_0)\right).$$

By assumption, $\frac{\partial F}{\partial y}(x_0, y_0)$ is invertible, so $\det(DG(x_0, y_0)) \neq 0$. Hence, $DG(x_0, y_0)$ is invertible.

By the Inverse Function Theorem, there exist open neighborhoods $W \subset \mathbb{R}^{n+m}$ around (x_0, y_0) and $Z \subset \mathbb{R}^{n+m}$ around $G(x_0, y_0)$, and a continuously differentiable map $H : Z \to W$ such that:

$$H \circ G = \mathrm{id}_W \quad \text{and} \quad G \circ H = \mathrm{id}_Z.$$

This implies that G is locally invertible near (x_0, y_0).

Now fix x in $\mathbb{R}^n$. Solving $G(x, y) = (x, 0)$ corresponds to:

$$\begin{bmatrix} x \\ F(x, y) \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}.$$

Thus, for each x near x_0, there exists a unique y near y_0 such that $F(x, y) = 0$. Define $g(x)$ to be this unique y.

Finally, since $g(x)$ is defined as part of the inverse map H, it inherits continuous differentiability from H. Therefore, $g(x)$ is a continuously differentiable function such that $F(x, g(x)) = 0$.

This proves the theorem. $\square$

Corollary 33

As a continuitation from the implict function theorem above, we now let $g : U \to V$ be that unique function such that $F(x, g(x)) = 0$ for $x \in U$. Then the derivative of $g(x)$, denoted $g'(x)$, can be expressed as:

$$g'(x) = -\left(\frac{\partial F}{\partial y}(x, g(x))\right)^{-1} \frac{\partial F}{\partial x}(x, g(x)),$$

where $\frac{\partial F}{\partial y}$ and $\frac{\partial F}{\partial x}$ are the Jacobian matrices of F with respect to y and x, respectively.

Proof. By the Implicit Function Theorem, the function $g : U \to V$ satisfies $F(x, g(x)) = 0$ for $x \in U$. Differentiating both sides with respect to x, we use the chain rule:

$$\frac{\partial}{\partial x}\left(F(x, g(x))\right) = 0.$$

Expanding the derivative:

$$\frac{\partial F}{\partial x}(x, g(x)) + \frac{\partial F}{\partial y}(x, g(x)) \cdot g'(x) = 0,$$

where:

- $\frac{\partial F}{\partial x}(x, g(x))$ is the $m \times n$ Jacobian matrix of F with respect to x,
- $\frac{\partial F}{\partial y}(x, g(x))$ is the $m \times m$ Jacobian matrix of F with respect to y,
- $g'(x)$ is the $m \times n$ Jacobian matrix of g with respect to x.

Rearranging, we solve for $g'(x)$:

$$\frac{\partial F}{\partial y}(x, g(x)) \cdot g'(x) = -\frac{\partial F}{\partial x}(x, g(x)).$$

Since $\frac{\partial F}{\partial y}(x, g(x))$ is invertible (by assumption), we multiply both sides by its inverse:

$$g'(x) = -\left(\frac{\partial F}{\partial y}(x, g(x))\right)^{-1} \frac{\partial F}{\partial x}(x, g(x)).$$

Thus, the derivative of $g(x)$ is given by the matrix expression:

$$g'(x) = -\left(\frac{\partial F}{\partial y}(x, g(x))\right)^{-1} \frac{\partial F}{\partial x}(x, g(x)).$$

$\square$

Let us try to inteprete the implicit function theorem

Fact 34 (Implicit function theorem intuition)

The implicit function theorem really just boils down to this: if I can write down m (sufficiently nice!) equations in $n + m$ variables, then, near any sufficiently nice solution point, there is a function of n variables which give me the remaining m coordinates of nearby solution points. In other words, I can, in principle, solve those equations and get the last m variables in terms of the first n variables. But (!) in general this function is only valid on some small set and won't give you all the solutions either.

Example 35
Here's a concrete example. Consider the equation $x^2 + y^2 = 1$. This is a single equation in two variables, and for a fixed $x_0 \neq \pm 1, y_0$ satisfying the equation, there is a function f of x such that $x^2 + f(x)^2 = 1$ for x near x_0, and $f(x_0) = y_0$. (Explicitly, for $y_0 > 0$, $f(x) = \sqrt{1-x^2}$, and for $y_0 < 0$, $f(x) = -\sqrt{1-x^2}$.) Notice that the function doesn't give you all the solution points — but this isn't surprising, since the solution locus of this equation is a circle, which isn't the graph of any function. Nonetheless, I have basically solved the equation and written y in terms of x.

2 Riemann Integral of Several Variables

2.1 existence of integral

Definition 36
A rectangle is a subset of Q of $\mathbb{R}^n$ of the form

$$Q = [a_1, b_1] \times \ldots \times [a_n, b_n]$$

where $a_i, b_i \in \mathbb{R}$. The **volume** of the rectangle is given by

$$v(Q) = (b_1 - a_1)\ldots(b_n - a_n)$$

The **width** of the rectangle is given by

$$\text{width}(Q) = \sup_i (b_i - a_i)$$

Reminder that *volume* is not the same as *measure*. The reinmann integral is defined using volume while the lebesgue intergral is defined using measure. They only conincide here because closed sets in $\mathbb{R}^n$ are measurable

Definition 37
A partition P of Q is an n-tuple $(P_1, \ldots, P_n)$ where each P_i is a partition of $[a_i, b_i]$

Definition 38
A partition P of $[a, b]$ is can be written as some $P = \{t_i : i = 1, \ldots, N\}$ where $t_1 = a < t_2 \ldots < t_N = b$. An interval belongs to P if and only if I is one of the intervals $[t_i, t_{i+1}]$. In the multivariable case, a rectangle $R = I_1 \times \ldots \times I_n$ belongs to a partition P of rectangle Q if for each i, the interval I_i belongs to P_i

Definition 39
Let $f : Q \to \mathbb{R}$ be a bounded function let P be partition of rectangle Q and R be a rectangle belong to P. We define

$$m_R f = \inf_R f \quad \text{and} \quad M_r f \sup_R f$$

and we define the upper and lower Rienman sums by

$$L(f,P) = \sum_R m_R(f)v(R) \quad \text{and} \quad U(f,p) = \sum_R M_r(f)v(R)$$

The upper and lower rieman integrals are defined by

$$\underline{\int_Q} f = \sup_P L(f,P) \quad h(x) = \overline{\int_Q} f = \inf_P U(f,P)$$

The riemann integral only exists

$$\underline{\int_Q} f = \overline{\int_Q} f$$

Note that we have implicity assumed

$$\int_{b_k}^{a_k} \dots \int_{b_1}^{a_1} f\, dx_1 \dots dx_k = \int_{I^k} f(\mathbf{x}) d\mathbf{x} \equiv \int_{I^k} f$$

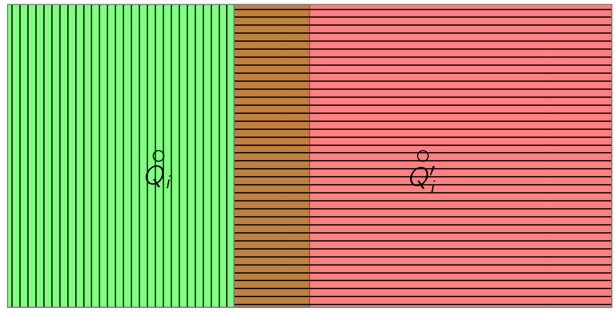
where $\mathbf{x} \in \mathbb{R}^k$ in case you wondering what happened to that d operator symbol which we always wrote in 1D. This integral will soon make sense after we learn about the **fubini theorem**

Theorem 40 (Lebesgue Criterion for Rienmann Integrability)
Let Q be a rectangle in $\mathbb{R}^n$ let $f : Q \to \mathbb{R}$ be a bounded function. Let D be the set of points of Q which fails to be continuous. Then $\int_Q f$ exists if and only if D has measure zero in $\mathbb{R}^n$

Proof. For the forward proof,since bounded we can choose M such that $|f(x)| \leq M$ for $x \in Q$. Now let D be covered by an open cover of rectangles $\{\operatorname{Int} Q_i\}$ whose total volume less than ε' since by definition of measure zero. Let elements in $Q - D$ be covered by an open cover of rectangles $\{\operatorname{Int} Q_i'\}$. Since continuous there exists for each element $a \in \operatorname{Int} Q_i'$ and ε

$$|f(x) - f(a)| \leq \varepsilon$$

a neighbourhood defined by $x \in Q_i' \cap Q$. Then the sets $\{Q_i\}$ and $\left\{Q_i'\right\}$ cover all of Q. Since Q is compact then this cover formed by these 2 sets are finite. Use the end points of each rectangle in $\{Q_i\}$ and $\{Q_i'\}$ to define a partition of P of Q that defines the subintervals(rectangles) $\{\mathcal{R}\}$. Let $\{R\}$ an $\{R'\}$ be sets of elements of $\{\mathcal{R}\}$ that lie *only* in Q_i or in Q_i' respectively



Then we have

$$\sum_R (M_R(f) - m_R(f))v(R) \leq 2M \sum_R v(R)$$

$$\sum_{R'} (M_{R'}(f) - m_{R'}(f))v(R') \leq 2\varepsilon' \sum_{R'} v(R')$$

so we have

$$\sum_R 2Mv(R) \leq \sum_{Q_i} 2Mv(Q_i) \leq 2M\varepsilon$$

and

$$\sum_{R'} 2\varepsilon' v(R') \leq 2\varepsilon' \sum_{\mathcal{R}} v(\mathcal{R}) = 2\varepsilon' v(Q)$$

therefore

$$U(f,P)-L(f,P) = \sum_R (M_R(f)-m_R(f))v(R) \leq 2M\sum_R v(R) + \sum_{R'} (M_{R'}(f)-m_{R'}(f))v(R') \leq 2\varepsilon' \sum_{R'} v(R') \leq 2M\varepsilon + 2\varepsilon' v(Q)$$

since Q is bounded $v(Q) < \infty$. To prove the other direction consider this lemma

Lemma 41

oscillations which we define as

$$\nu(f,a) = \inf_{\delta>0} [M_\delta(f) - m_\delta(f)]$$

for a neighbourhood of radius $\delta > 0$ around $a \in Q$.

Then $\nu(f,a) = 0$ if and only if f is continuous at a.

Proof. If continuous for any $\varepsilon > 0$ there exists a neighbourhood $|x - a| < \delta$ where $|f(x) - f(a)| < \varepsilon$. Then since in this neighbourhood $f(a)$ is related to $M_\delta(f)$ and $m_\delta(f)$ like so by definition

$$f(a) \leq M_\delta(f) < f(a) + \varepsilon \quad \text{and} \quad f(a) \geq m_\delta(f) > f(a) - \varepsilon$$

from which we can obtain

$$M_\delta(f) - m_\delta(f) < \varepsilon$$

so because ε is abrituary $\nu(f,a) = 0$ Conversely if $\nu(f,a) = 0$ by definition of inf there exists

$$M_\delta(f) - m_\delta(f) \geq \nu(f,a) > M_\delta(f) - m_\delta(f) - \varepsilon$$

so

$$M_\delta(f) - m_\delta(f) < \varepsilon$$

Therefore any x in this neighbouhood $|x-a|<\delta$ we have

$$m_\delta(f) \le f(x) \le M_\delta(f)$$

it follows $|f(x)-f(a)|<\varepsilon$ since the oscillation is the greatest possible difference in this neighbourhood and even it is less than ε. ☐

Now back to proving the other direction consider

$$D_m = \left\{ a|\nu(f,a) \ge \frac{1}{m} \right\}$$

Suppose f is integrable that is

$$U(f,P) - L(f,P) < \varepsilon$$

We know from the above lemma must exists in one of the sets D_m. Now consider points of any D_m that in Bd R of an R defined by the partion and points that in the interior of an R instead. Since BdR has measure zero we already settled that case. Now for the interior cases, we will show that these rectangles whose interiors contain points in D_m are of measure zero. First consider that since $a \notin \text{Bd}\, R$ there exists a radius $\delta > 0$ around it contained in R in which case it will define an oscillation such that a belongs to some D_m. In such cases we have

$$\frac{1}{m} \le \nu(f,a) \le M_\delta(f) - m_\delta(f) \le M_\mathcal{R}(f) - m_\mathcal{R}(f)$$

Now consider the sum over all rectangles R_i that contain points in this fixed D_m and assume there are k of them

$$\sum_{i=1}^{k} \left(\frac{1}{m}\right) v(R_i) \le \sum_{\mathcal{R}} \left(\frac{1}{m}\right) v(\mathcal{R}) \le \sum_{\mathcal{R}} (M_\mathcal{R}(f) - m_\mathcal{R}(f)) v(\mathcal{R}) = U(f,P) - L(f,P) < \varepsilon$$

Thus the set of discontinuous points are of measure zero. ☐

Theorem 42

Let Q be a rectangle in $\mathbb{R}^n$. Let $f: Q \to \mathbb{R}$. Assume that f is reinmann integrable over Q.

1. If f vanishes($f=0$) except on a set of zero then $\int_Q f = 0$
2. If f is non-negative and if $\int_Q f = 0$ then f vanishes except on a set of measure zero

Proof. For(1) Suppose f vanishes except on a set E of measure zero. We know from the definition of exterior measure in `functional analysis notes` it is not possible for any rectangle to be contained in E. Therefore for any partition we know that E must within the rectangle defined by it. Therefore every rectangle must must contain a point not in E where $f(x)=0, x \in \forall R$. Therefore we have

$$L(f,P) \le 0 \quad \text{and} \quad U(f,p) \ge 0$$

Since these inequalities hold for all P we have

$$\underline{\int}_Q f \le 0 \quad \text{and} \quad \overline{\int}_Q f \ge 0$$

And if $\int_Q f$ exists(that is f is integrable), then clearly $\int_Q f = 0$ For (2) suppose $f(x) \ge 0$ and $\int_Q f = 0$. Then, from the previous theorem we know that the set of discontinuous points is one such set of measure zero.Hence we can prove by contradiction by supposing that f does not vanish on the set of continuous points. That is let f be continuous at a and that $f(a) > 0$. Then there exists for every $\varepsilon > 0$

$$-\frac{\varepsilon}{2} < f(x) - f(a) < \frac{\varepsilon}{2} \quad \text{for some} \quad |x-a| < \delta$$

Let $f(a) = \varepsilon$ so we have $\frac{\varepsilon}{2} < f(x) < 1.5\varepsilon$. It follows that for the partition that has a rectangle R_0 corresponding to this neighbourhood of δ around a, we have $m_{R_0}(f) \geq \varepsilon/2$ while $m_R(f) \geq 0$ for all R(since no negative function). Then we have

$$L(f,P) = \sum_R m_R(f)v(R) \geq (\varepsilon/2)v(R_0) > 0$$

but $L(f,P) \leq \int_Q f = 0$ so we have a contradiction. $\square$

For the rest of this set of notes we are using the *Rienmann Integral*

2.2 fubini theorem

Theorem 43 (Fubini)
Let $f : Q \to \mathbb{R}$ be a bounded function and $Q = A \times B$ a rectangle as defined above. We write $f = f(x,y)$ where $x \in A$ and $y \in B$. Fixing $x \in A$ we can define a function $f_x : B \to \mathbb{R}$ by $f_x(y) = f(x,y)$. Since this function is bounded we can define new functions $g, h : A \to \mathbb{R}$ by

$$g(x) = \underline{\int}_B f_x \quad h(x) = \overline{\int}_B f_x$$

Then fubini theorem states that if f is integrable over Q then g, h are integrable A so we have

$$\int_A \int_B f_x = \int_Q f$$

So far we have been considering reinmann integrals over rectangles. We now extend it to more general regions

Definition 44 (Integration over general regions)
Let S be a bounded set in $\mathbb{R}^n$ and let $f : S \to \mathbb{R}$ be a bounded function. Let $f_S : \mathbb{R}^n \to \mathbb{R}$ be a map defined by

$$f_S(x) = \begin{cases} f(x) & x \in S \\ 0 & x \notin S \end{cases}$$

Let Q be a rectangle such that the interior of $Q \supset \overline{S}$ The map f is riemann integrable over S if f_S is riemann integrable over Q that is

$$\int_S f = \int_Q f_s$$

Theorem 45
Let Q be a rectangle such that Int $Q \supset \overline{S}$. The characteristic function defined by

$$1_S(x) = \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases}$$

Then

$$\int_Q 1_S = \int_S 1$$

if and only if Bd S is of measure zero

Proof. recall 21 that both Int S an Ext S are open sets therefore $x \in Q$ which are in these sets correspond to when $1_S = 1$ and $1_S = 0$. Obviously continuous just consider $|1_S(x) - 1_S(x)| < \varepsilon \to |x - y| < \delta$. However if in Bd S every neighbourhood of x ther exists point in Ext S where $1_S = 0$ and Int S where $1_S = 1$ so obviously discontinuous. Therefore by the conclusion follows.

Definition 46
The set S is **rectifiable** if the boundary of S is of measure zero. If S is rectifiable then

$$v(S) = \int_S 1$$

This makes sense because by definition we calculate volume by using only rectangles so we have to approximate using the integral of some characteristic function of S over rectangles. However from above we know that this is not possible if Bd S does not have measure zero because these are points that will be discontinuous on our characteristic function. Over here we did not require f to be compact because 1 function is bounded by definition.

Lemma 47
Rectangles are rectifiable

The boundary points of a closed rectangle

$$Q = [a_1, b_1] \times \ldots \times [a_n, b_n]$$

correspond to faces of Q which is the union of all subset of Q in which $x_i = a_i$ or $x_i = b_i$. These faces are of measure zero because each of these subsets of Q can be covered by a rectangle like for example

$$\text{face } \mathrm{Q}_{x_i = a_i} = [a_1, b_1] \times \ldots [a_i, a_i + \delta] \ldots \times [a_n, b_n]$$

where δ can be made arbitrarily small. Therefore the volume of such faces can be made arbitrarily small as well.

Lemma 48 (ε neighbourhood theorem)
Let X be a compact subspace of $\mathbb{R}^n$ and let U be an open set in $\mathbb{R}^n$ containing X. Then there is an $\varepsilon > 0$ such that the ε neighbourhood of X is containd in U

proof Let C be a fixes subset of $\mathbb{R}^n$ for each $x \in \mathbb{R}^n$ define

$$d(x, C) = \inf \{|x - c| \,|c \in C\}$$

We show that the distance from x to C is a continuous function of x

$$d(x, C) - |x - y| \leq |x - c| - |x - y| \leq |y - c|$$

by triangle inquality and since this applies for all $c \in C$ we have

$$d(x, C) - |x - y| \leq d(y, C)$$

so we have

$$d(x, C) - d(y, C) \leq |x - y|$$

Theorem 49

Let A be an open subset of $\mathbb{R}^n$. There exists a sequence of compact rectifiable sets C_N, $N = 1, 2, 3 \ldots$ such that

$$C_N \subseteq \text{Int } C_{N+1}$$

and

$$\bigcup C_N = A$$

The set $\{C_N\}$ is called the **exhaustion** of A

Proof. Take the complement of A namely $B = \mathbb{R}^n - A$ Define

$$D_N = \left\{x \in A : d(x, B) \geq \frac{1}{N} \text{ and } |x| \leq N\right\}$$

Since A is open there exists a ε where $\{y : |y - x| \leq \varepsilon\}$ contained in A. Therefore $d(x, B) \geq \varepsilon$. Let $\frac{1}{N} < \varepsilon$ so $d(x, B) \geq \frac{1}{N}$ which is certainly possible by archimedian property. Now we select all such x contained in the neighbourhood $\{x : |x - 0| \leq N\}$. Because for any $x \in A$, it has to be in at least one of such D_N but there certainly exists a finite distance between any point in x and 0 and that $\lim_{n\to\infty} \frac{1}{n} < \varepsilon$. Therefore

$$\bigcup D_N = A$$

It is easy to check that $D_N \subseteq \text{Int } D_{N+1}$ just look at $|x| \leq N$ where N is the boundary so

$$\text{Int } D_N = \{x \in D_N \, |x| < N\}$$

To ensure our sets are rectifiable we construct an open cover of D_N by

$$\{\text{Int } Q_j : p \in D_N \text{ and } Q_j \subseteq \text{Int } D_{N+1}\}$$

where Q_j is a rectangle. Because D_N is compact there is finite subcover where

$$D_N \subset \text{Int } Q_{p_1} \cup \ldots \cup \text{Int } Q_{p_r} = \text{Int } C_N$$

Taking the closure of this subcover which we denote to be

$$C_N = Q_{p_1} \cup \ldots \cup Q_{p_r}$$

So we see that

$$\ldots \text{Int } D_N \subset D_N \subset \text{Int } C_N \subset C_N \subset \text{Int } D_{N+1} \ldots$$

Thus the sequence of sets $C_N \subseteq \text{Int } C_{N+1}$ is rectifiable recall the lemma above $\square$

So far we have generalized our integral over rectangles to compact and rectifiable sets. We now generalize it to open sets too by definining the improper integral.

Definition 50

Let A be an open set in $\mathbb{R}^n$ and let $f : A \to \mathbb{R}$ be a continuous function.For the moment we assume that $f \geq 0$. Let $D \subseteq A$ be a compact and rectifiable set. Then $f|D$ is bounded so $\int_D f$ is well-defined as we can get real sup/inf for any partition of our choice and because rectifiable points of discontinuity measure zero. Now consider the set of such integrals denoted by

$$\# = \left\{ \int_D f : D \subseteq A, D \text{ compact and rectifiable} \right\}$$

Then **improper integral** of f over A exists if $*$ is bounded and is then defined by

$$\int_A^{\#} f \equiv \sup \int_D f = \text{ improper intergral of f over A}$$

Remark 51. *First of all we are obviously assuming a continuous function if not every point will be a point of discontinuity. Note that the reason the set the integral is being defined on is compact is really only to ensure the continuous function is bounded. If we have say S is a rectifiable set so*

$$\int_S 1 = \int_{\text{Int } S} 1 = v(\text{Int } S) = v(S)$$

this still make sense even through $\text{Int } S$ is open because the 1 function is clearly bounded so the upper and lower rieman sums are defined. However in any case if S is not rectifiable, then this integral won't make sense recall 46

Theorem 52

Suppose improper integral $\int_A f$ exists then

$$\int_A f = \lim_{n \to \infty} \int_{C_n} f$$

where $\{C_n\}$ is a sequence of compact rectifiable sets whose union is A

Proof. We assume the existence of $\cup C_n = A$ by the previous theorem. First by monotonicity the sequence $\left\{ \int_{C_n} f \right\}$ is increasing so it converges if and only if it is bounded. Suppose that f is integrable over A and let D range over all compact rectifiable subsets of A. Then clearly

$$\int_{C_n} f \leq \sup_D \left\{ \int_D f \right\} = \int_A f$$

Therefore $\left\{ \int_{C_n} f \right\}$ is bounded. Hence

$$\lim_{n \to \infty} \int_{C_n} f \leq \int_A f$$

To prove the other direction since

$$\text{Int } C_1 \subset \text{Int } C_2 \ldots$$

(that is they are circles embedded in onr another) and

$$\cup C_n = A, D \subset A$$

(since D compact is a finite union of C_n covers D), there exists some Int C_M where

$$\int_D f \leq \int_{C_m} f \leq \lim_{n\to\infty} \int_{C_n} f$$

since $D \subset \text{Int } C_m$ and $f \geq 0$. Because D is arbituary we may take the suprenum of the LHS Therefore

$$\int_A f = \lim_{n\to\infty} \int_{C_n} f$$

Theorem 53

Let A be a bounded open set in $\mathbb{R}^n$. Let $f : A \to \mathbb{R}$ be a bounded continuous function. Then the extende integral $\int_A f$ eixists. If the orderinary integral $\int_A f$ also exists then these two integrals are equal.

Proof. Let Q be a rectangle containing A. We first show that the extended integral of f exists. Choose M so that $|f(x)| \leq M, \forall x \in A$. Then for any compact recitifbale subset D of A we have

$$\int_D |f| \leq \int_D M \leq Mv(Q)$$

Now consider the case where f is non-negative. Suppose the ordinary integral $\int_A f$ exists. Then since $f = f_A$ on D

$$\int_D f = \int_D f_A \leq \int_Q f_A = \int_A f$$

to prove the other direction like how we did previously let $D = R_1 \cup \ldots \cup R_k$ where these are subrectangles of rectangles R defined by partition P of Q that lie in A. Then

$$L(f_A, P) = \sum_{i=1}^{k} m_{R_i}(f)v(R_i)$$

since $m_R(f_A) = m_R(f)$ if R is contained(not just intersect) in A but $m_R(f_A)$ otherwise. Then we can relate it to

$$\sum_{i=1}^{k} m_{R_i}(f)v(R_i) \leq \sum_{i=1}^{k} \int_{R_i} f = \int_D f \leq \int_A f$$

where equalities follow by compariions,additivity and definition of the extended integral respectively.

2.3 support and compact support

Definition 54

Let U be an open set in $\mathbb{R}^n$ and let $f : U \to \mathbb{R}$ be a continuous function. The **support** of f is

$$\text{supp } w = \overline{\{x \in U : f(x) \neq 0\}}$$

The function f is **compactly supported** if supp w is compact. We define

$$\mathcal{C}_0^k(U) = \text{the set of compactly supported } \mathcal{C}^k \text{ functions on } U$$

That is to say if $x \in (\text{supp } w)^c$ which is an open set then there is a neighbourhood of x on which the function w vanishes identically.

2.4 partitions of unity

We now consider an alternative way to define an improper integral. In the above we defined the **exhaustion** of an open set A in which we can use the define the improper integral like so

$$\int_A f = \lim_{n\to\infty} \int_{C_n} f$$

In our new approach we instead breaking up the integration region into a sequence of compact rectifiable sets now we break the up the function into a sequence of functions we call the **partition of unity**

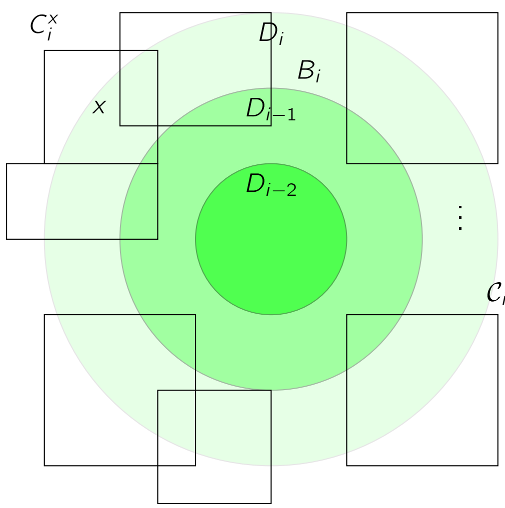
Lemma 55

Let $\mathcal{A}$ be a collection of open sets in $\mathbb{R}^n$. Let A be their union. There exists a countable collection $Q_1, Q_2 \dots$ of rectangles contained in A such that

1. the sets $\operatorname{Int} Q_i$ cover A
2. Each Q_i is contained in an element of $\mathcal{A}$
3. Each point of A has a neighbourhood that intersects only finitely many of the sets Q_i

Proof. Let $D_1, D_2 \dots$ be a sequence of compact subsets of A whose union is A such that $D_i \subset \operatorname{Int} D_{i+1}$. We know from previously this is called the exhaustion of A and it exists. For each i define

$$B_i = D_i - \operatorname{Int} D_{i-1}$$



For each $x \in B_i$ choose a closed cube C_x^i centered at x that are disjoint from D_{i-2}(that is to ensure there is some space to define a cube. Because an open. Therefore the interiors of each cube C_x^i cover B_i. Because B_i is compact there exists a finite cover using such cubes which we denote as $\mathcal{C}_i$ corresponding to B_i. Let $\mathcal{C} = \bigcup_i \mathcal{C}_i$. The sets $\operatorname{Int} C_x$ cover A as clearly every x must lie in some B_i and $\operatorname{Int} C_i$ covers B_i so x must be in some C_x^i, proving property (1). We can simply make each C_x^i small enough such that they are contained in an element of $\mathcal{A}$. This is possible because even in the worst case(union of smallest open sets) of $A = \bigcup \mathcal{A}_i$, each $\mathcal{A}_i$ by definition of open set must still have a neighbourhood($r > 0$) and therefore at very least lie in some B_i. In which case the covering by C_x takes care of it. In that case in which we have property (2) Finally because C_i is a finite covering, every open set $\operatorname{Int} D_i$ can only intersect a finite number of cubes corresponding to those from $C_1, \dots, C_{i+1}$. Therefore since every point $x \in A$ which must obviously lie in some $\operatorname{Int} D_i$ which is precisely an open neighbourhood that intersects with only finitely many sets C_x^i.

So (3) is proven. Finally let $\{Q_i\}$ be the sequence of rectangles formed from enumerating through all C_x^i(countable union of countable sets are countable)

Theorem 56 (Partitions of Unity)
Let $\mathcal{A}$ be a collection of open sets in $\mathbb{R}^n$. Let A be their union. There exist a sequence $\{\phi_i\}$ of continuous functions $\phi_i : \mathbb{R}^n \to \mathbb{R}$ known as the **partion of unity** such that

1. $\phi_i(x) \geq 0$ for all x
2. The set $S_i = \operatorname{supp} \phi_i$ is contained in A
3. Each point of A has a neighbourhood that intersects only finitely many of the sets S_i
4. $\sum_{i=1}^{\infty} \phi_i(x) = 1$ for all $x \in A$
5. The functions ϕ_i are of the class $\mathcal{C}^\infty$
6. The sets S_i are compact
7. For each i the set S_i is contained in an element of $\mathcal{A}$ in which case we say $\{\phi\}$ is a partition of unity **subordinate** to $\mathcal{A}$

Proof. Take a sequence of rectangles $\{Q_i\}$ that satisfy the above lemma. Then for each i *define* $\psi_i : \mathbb{R}^n \to \mathbb{R}$ be a $\mathcal{C}^\infty$ function. In particular if we use a class of functions known as **bump functions** which are smooth.Then we have satisfy (5). These bump functions are defined to be positive on $\operatorname{Int} Q_i$ and zero elsewhere.This means $\phi_i(x) \geq 0$, since by construction ϕ_i can only be zero or bigger than zero. In that case we have (1). Furthermore $\operatorname{supp} \phi_i = Q_i$ since Q_i is obviously the closure of $\operatorname{Int} Q_i$. $\sup \phi_i$ being closed in $\mathbb{R}^n$ is therefore compact proving (6). (7),(3),(2) can be seen by the previous lemma. Therefore this implies the series

$$\lambda(x) = \sum_{i=1}^{\infty} \phi_i(x)$$

for all $x \in A$ converges because only a finite number of $\phi_i(x)$ in this sum is greater than zero. Therefore we can prove (4) by defining

$$\phi_i(x)_= \psi_i(x)/\lambda(x)$$

which obviously "normalizes" $\sum \phi_(x)$

Example 57
The bump function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} e^{-\frac{1}{x}} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Then let $g(x) = f(x)f(1-x)$(this defines our interval of support). Clearly g is positive for $0 < x < 1$ and is of class C^∞(exponetial function)

Theorem 58
Let A be open in $\mathbb{R}^n$ and let $f : A \to \mathbb{R}$ be continuous. If f vanishes outside the *compact* subset C of A then the integrals $\int_A f$ and $\int_C f$ exist and are equal.

Proof. The integral $\int_C f$ exists because C is bounded by 53. Let C_i be a sequence of compact rectifiable sets whose uunion is A and that $C_i \subseteq \operatorname{Int} C_{i+1}$ for each i. Since C is compact it is covered by finitely many sets of $\operatorname{Int} C_i$. So there exists $\operatorname{Int} C_m$ where

$$\int_C f = \int_{C_N} f$$

since f vanishes outside C for $N \geq M$. This shows that $\lim_{n\to\infty} \int_{C_n} f$ exists as well and equals $\int_A f$ as $A = \bigcup C_i$

Theorem 59

Let A be open in $\mathbb{R}^n$ and let $f : A \to \mathbb{R}$ be continuous. Let $\{\phi_i\}$ be a partition of unity on A having *compact supports*. The integral $\int_A f$ exists if and only if the series

$$\sum_{i=1}^{\infty} \int_A \phi_i \, |f|$$

converges. In that case then

$$\int_A f = \sum_{i=1}^{\infty} \int_A \phi_i f$$

Proof. Denote $\operatorname{supp} \phi_i$ by S_i.Let D be a compact rectifiable subset of A that is we have a well defined $\int_D f$. Because D is bounded, is surely covered by finite subcover of $\operatorname{supp} \phi_i$. Hence there exists an M such that for all $i > M$ the function ϕ_i vanishes outside D. That is

$$f(x) = \sum_{i=1}^{M} \phi_i(x) f(x) = \sum_{i=1}^{\infty} \phi_i(x) f(x)$$

where $\sum_{i=1}^{M} \phi_i(x) = \sum_{i=1}^{\infty} \phi_i(x) = 1$ recall this is possible from 56. Thus

$$\int_D f = \sum_{i=1}^{M} \int_D \phi_i f \leq \sum_{i=1}^{M} \int_{D \cup S_i} \phi_i f = \int_A \phi_i f \leq \sum_{i=1}^{\infty} \int_A \phi_i f$$

the first second equality follows by monotoncity and the third equality follows by the previous lemma because $\phi_i f$ vanishes outside $D \cup S_i \subset A$. Since the RHS is bounded we know the extended integral exists. To see this just take sup over D on both sides we have

$$\int_A f \leq \sum_{i=1}^{\infty} \int_A \phi_i f$$

To prove the other direction again by the fact let $D = S_1 \cup \ldots \cup S_M$ we have by the previous lemma

$$\sum_{i=1}^{M} \int_A \phi_i f = \sum_{i=1}^{M} \int_D \phi_i f = \int_D f \leq \int_A f$$

where the last equality is by definition of the extended integral. Now because partial sums are bounded taking the limits on the LHS we have

$$\sum_{i=1}^{\infty} \int_A \phi_i f \leq \int_A f$$

so the conclusion

$$\int_A f = \sum_{i=1}^{\infty} \int_A \phi_i f$$

follows

Corollary 60
There exists a function $p \in \mathcal{C}^\infty(\mathbb{R})$ where $\int_A p = 1$ where f is compactly supported on A.

Proof. Note of all the Lebesgue and Riemann integral agree here.
Let $p = \sum \phi_i$ which is the sum of the partitions of unity. Then

$$\begin{aligned}\int_A p d\mu &= \int_A \sum \phi_i d\mu \\ &= \int_A 1 d\mu \\ &= m(A) < \infty\end{aligned}$$

Since A is bounded in $\mathbb{R}$ its measure is certainly finite.Or in Riemann integral terminology the volume $v(A) < \infty$ So we may define $p(x) = \frac{\sum \phi_i(x)}{m(A)}$ since we know the partitions of unity exists for compactly supported functions.

3 multi-linear algebra

Theorem 61 (Dual Space)
Let the **dual space** V^* be the set of all linear functions $\ell : V \to \mathbb{R}$ We claim that the basis functions of V^* are

$$e_i^*(v) = \begin{cases} e_i^*(e_j) = 1 & i = j \\ e_i^*(e_j) = 0 & i \neq j \end{cases}$$

and that $\dim V = \dim V^*$

Remark 62. *This makes sense. Think of each $e_i^*(v)$ to be a basis function. It is in charge of identifying e_i in V and assigning a scalar multiplier to it like so $\lambda_i e_i^*(e_i) = \lambda_i$ which then defines the map $\ell : V \to \mathbb{R}$. So the basis is $\{e_i^*(v)\}$ while the coeffients in its span are $\{\lambda_i\}$*

Proof. Let $\dim V = n$ Any ℓ can represented by

$$\ell(v) = \ell(\sum_i^n c_i e_i) = \sum_i^n c_i \ell(e_i) = \sum_i^n c_i \lambda_i e_i^*(e_i) \in \mathbb{R}$$

where $\lambda_i, c_i \in \mathbb{R}$ and $\ell(e_i) = \lambda_i$. Therefore $\{e_i^*(v)\}$ spans V^* so write the same thing more compactly

$$\ell(v) = \sum_i^n \lambda_i e_i^*(v)$$

Because when ever $e_i \in v$ then $e_i^*(v)$ will be 1 so the sum is the same. We don't have to write c_i because it will appear by itself by linearity, it is not dependent on the function. We now need to prove linear independence. Suppose $\ell = 0$ is the zero function. Then for every e_i we must have

$$0 = \ell(e_i) = \lambda_i$$

Therefore it follows that whenever a linear operator is the zero function then all $\lambda_i = 0$ thus the below is the trivial sum

$$\ell(v) = \sum_i^n \lambda_i e_i^*(v) = 0$$

It clearly follows that for every $\ell \in V^*$ there it is represented with unqiue set of λ_i and that $\dim V = \dim V^* = n$ $\square$

3.1 tensors

Definition 63

Let V be an n-dimensional vector space. Suppose

$$V^k = \underbrace{V \times \ldots \times V}_{\text{k times}}$$

The function $f : V^k \to \mathbb{R}$ is **linear in its ith variable** if given fixed vectors v_j for $j \neq i$ the map $T : V \to \mathbb{R}$ defined by

$$T(\mathbf{v}) = f(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k)$$

is linear The function f is said to **multilinear** if it is linear in the ith variable for each i. Such a function is also called a **k-tensor** or a **tensor of order k** on V. The set of all k-tensors are deonted by $\mathcal{L}^k(V)$

Example 64

Suppose $T \in \mathcal{L}^2$ then

$$\begin{aligned} T(\lambda_{1,1}v_{1,1} + \lambda_{1,2}v_{1,2}, \lambda_{2,1}v_{2,1} + \lambda_{2,2}v_{2,2}) &= \lambda_{2,1}T(\lambda_{1,1}v_{1,1} + \lambda_{1,2}v_{1,2}, v_{2,1}) + \lambda_{2,2}T(\lambda_{1,1}v_{1,1} + \lambda_{1,2}v_{1,2}, v_{2,2}) & (1) \\ &= \lambda_{1,1}\lambda_{2,1}T(v_{1,1}, v_{2,1}) + \lambda_{1,1}\lambda_{2,2}T(v_{1,1}, v_{2,2}) & (2) \\ &\quad + \lambda_{1,2}\lambda_{2,1}T(v_{1,2}, v_{2,1}) + \lambda_{1,2}\lambda_{2,2}T(v_{1,2}, v_{2,2}) & (3) \end{aligned}$$

We apply linearity in the second factor in (1). Then for each term in (1) we apply linearity in the 1st factor to obtain (2) and (3) respectively.Hence we have in summary

$$T(\lambda_{1,1}v_{1,1} + \lambda_{1,2}v_{1,2}, \lambda_{2,1}v_{2,1} + \lambda_{2,2}v_{2,2}) = \sum_{i=1}^2 \sum_{j=1}^2 \lambda_{1,i}\lambda_{2,j}T(v_{1,i}, v_{2,j})$$

Example 65

the set $\mathcal{L}^1$ is the set of all linear transformations $f : V \to \mathbb{R}$. It is sometimes called the **dual space** of V and denoted by V^*

Example 66

$\mathcal{L}^0 = \mathbb{R}$

Definition 67 (Tensor Product)
Let $T_i \in \mathcal{L}^{k_i}, i = 1, 2$ define $k = k_1 + k_2$. we define the **tesor product** of T_1 and T_2 to be $T_1 \otimes T_2 : V_k \to \mathbb{R}$ defined by

$$T_1 \otimes T_2(v_1, \dots . v_k) = T_1(v_1, \dots, v_{k_1})T_2(v_{k_1+1}, \dots, v_k)$$

We can easily see that for each $v_i, i, \dots, k$ linearity still applies. So we can conclude $T_1 \otimes T_2 \in \mathcal{L}^k$

Proposition 68
Tensor products satisfy

- (Associative Law) $(T_1 \otimes T_2) \otimes T_3 = T_1 \otimes (T_2 \otimes T_3)$
- (Left and right distributive laws)Suppose $T_i \in \mathcal{L}^{k_i}, i = 1, 2, 3$ and assume $k_1 = k_2$ then

$$(T_1 + T_2) \otimes T_3 = T_1 \otimes T_3 + T_2 \otimes T_3 \quad \text{and} \quad T_3 \otimes (T_1 + T_2) = T_3 \otimes T_1 + T_3 \otimes T_2$$

- Non commutative

Proof. Associative law is pretty obvious from how tensor product is defined. It is also clear that tensor products are not commutative.Consider

$$\begin{aligned} T_1 \otimes T_2(v_1, \dots . v_k) &= T_1(v_1, \dots, v_{k_1})T_2(v_{k_1+1}, \dots, v_k) \\ &\neq T_2(v_{k_1+1}, \dots, v_k)T_1(v_1, \dots, v_{k_1}) \\ &= T_2 \otimes T_1(v_{k_1+1}, \dots, v_k, v_1, \dots, v_{k_1}) \end{aligned}$$

However distributive properties mainly reflect linearity. Consider the following tensor product in $\mathcal{L}^2$

$$(T_1 \otimes (T_2 + T_3))(x, y) = T_1(x)(T_2 + T_3)(y) = T_1(x)(T_2(y) + T_3(y)) = T_1(x)T_2(y) + T_1(x)T_3(y) = T_1 \otimes T_2 + T_1 \otimes T_3$$

Definition 69
If a tensor can be defined by

$$T = \ell_1 \otimes \dots \otimes \ell_k$$

where each $\ell_1 \in \mathcal{L}^1$ that is we can also write

$$T(v_1, \dots, v_k) = \ell_1(v_1) \dots \ell_k(v_k)$$

then we say this is a **decomposable k-tensor**

Theorem 70

A multi index I of length k is a set of integers $(i_r, \dots, i_k)$ such that every element $i_r, r \in [1, k]$ is in the range $1 \le i_r \le n$ Suppose we have

$$e_i^*(v) = \begin{cases} e_i^*(e_j) = 1 & i = j \\ e_i^*(e_j) = 0 & i \neq j \end{cases}$$

Define a decomposable k-tensor $e_I^* = e_{i_1}^* \otimes \dots \otimes e_{i_k}^* \in \mathcal{L}^k$. That is we write

$$e_I^*(e_{i_1}, \dots, e_{i_k}) = e_{i_1}^*(e_{i_1}) \dots e_{i_k}^*(e_{i_k})$$

therefore suppose $J = (j_1, \dots, j_k)$ be a multi-index of lenghth k. Then

$$e_I^*(e_{j_1}, \dots, e_{j_k}) = e_{i_1}^*(e_{j_1}) \dots e_{i_k}^*(e_{j_k}) = \begin{cases} e_i^*(e_j) = 1 & I = J \\ e_i^*(e_j) = 0 & I \neq J \end{cases}$$

We claim that this is in fact a basis for $\mathcal{L}^k$ and that $(\dim V)^k = \dim \mathcal{L}^k$

Proof. Let $\dim V = n$. Any $\ell \in \mathcal{L}^k$ can be represented by

$$\begin{aligned} \ell(\underbrace{v, \dots, v}_{\text{k-times}}) &= \ell(\sum_{i_1}^{n} c_{i_1} e_{i_1}, \dots, \sum_{i_k}^{n} c_{i_k} e_{i_k}) && (1) \\ &= \sum_{i_1}^{n} c_{i_1} \dots \sum_{i_k}^{n} c_{i_k} \ell(e_{i_1}, \dots, e_{i_k}) && (2) \\ &= c_I \sum_{I} e_I^*(e_{i_1}, \dots, e_{i_k}) \lambda_I && (3) \end{aligned}$$

Where $\lambda_I, c_I \in \mathbb{R}$. Like previously we can write this same thing more compactly as

$$\ell(\underbrace{v, \dots, v}_{\text{k-times}}) = \sum_{I} e_I^*(\underbrace{v, \dots, v}_{\text{k-times}}) \lambda_I$$

(2) follows by linearity. Therefore we have shown that $\{e_I^*(e_{i_1}, \dots, e_{i_k})\}$ spans $\mathcal{L}^k$. We now show linear independence. Suppose $\ell(\underbrace{v, \dots, v}_{\text{k-times}}) = 0$ be the zero function. Then for every $\ell(e_{i_1}, \dots, e_{i_k})$ corresponding to a permutation of I we must have $\lambda_I = 0$ because $\ell(e_{i_1}, \dots, e_{i_k}) = \lambda_I$. Therefore the only result in (3) where $\ell = 0$ is the trivial sum thus proving linear independence. Therefore the dimension of $\mathcal{L}^k$ is n^k which is the number of permutations of I.

3.2 dual transformation $\mathcal{L}^k(W) \to \mathcal{L}^k(V)$

Definition 71 (Dual Transformation)

Let V, W be vector spaces and let $A : V \to W$ be a linear map. Let $T \in \mathcal{L}^k(W)$ and define the new map $A^*T \in \mathcal{L}^k(V)$. We deonte

$$A^*T(v_1, \dots, v_k) = T(Av_1, \dots, Av_k)$$

Theorem 72

Let $T : V \to W$ be a linear transformation. Let

$$T^* : \mathcal{L}^k(W) \to \mathcal{L}^k(V)$$

be a dual transformation then

1. T^* is linear
2. $T^*(f \otimes g) = T^*f \otimes T^*g$
3. if $S : W \to X$ is a linear transformation then $(S \circ T)^*f = T^*(S^*f)$

Proof. For (1) Consider

$$\begin{aligned}(T^*(af+bg))(v_1,\dots,v_k) &= (af+bg)(T(v_1),\dots,T(v_k)) \\ &= af(T(v_1),\dots,T(v_k)) + bg(T(v_1),\dots,T(v_k)) \\ &= aT^*f(v_1,\dots,v_k) + bT^*f(v_1,\dots,v_k)\end{aligned}$$

For (2) consider

$$\begin{aligned}T^*(f \otimes g)(v_1,v_2) &= (f \otimes g)(Tv_1,Tv_2) \\ &= f(Tv_1)g(Tv_2) \\ &= T^*f(v_1) + T^*g(v_2)\end{aligned}$$

For (3) consider the following commutative diagrams

$$\begin{tikzcd} & W \arrow[dr, "S"] & \\ V \arrow[ur, "S"] \arrow[rr, "S\circ T"] & & X \end{tikzcd} \qquad \begin{tikzcd} & \mathcal{L}^k(W) \arrow[dr, "T^*"] & \\ \mathcal{L}^k(X) \arrow[ur, "S^*"] \arrow[rr, "(S\circ T)^*"] & & \mathcal{L}^k(V) \end{tikzcd}$$

Definition 73

A **permutation** of order k is a bijective map

$$\sigma : \{1,\dots,k\} \to \{1,\dots,k\}$$

and that σ^{-1} exists. Given two permutations σ_1, σ_2 we can construct the composite permutation as

$$\sigma_1\sigma_2 = \sigma_1 \circ \sigma_2(i) = \sigma_1(\sigma_2(i))$$

The **symmetric group** is denoted by S_k. Recall from `artin algebra I`, S_k has $k!$ elements, each being a lenght k set that contains *positional indices* not values!

Consider the following special types of permutations called transpositions we are basically 2-cycle permutation groups

Example 74

The permuation $\tau_{i,j}$ defined by

$$\begin{aligned}\tau_{i,j}(i) &= j\\ \tau_{i,j}(j) &= i\\ \tau_{i,j}(\ell) &= \ell \quad , \quad \ell \neq i,j\end{aligned}$$

is called a **transposition**. In cyclic notation(recall `artin algebra I`) we denote

$$\tau_{i,j} = (i,j)$$

Example 75

The permutation $\tau_{i,i+1}$ defined by

$$\begin{aligned}\tau_{i,i+1}(i) &= i+1\\ \tau_{i,i+1}(i+1) &= i\\ \tau_{i,i+1}(j) &= j \quad , \quad j \neq i, i+1\end{aligned}$$

is called an **elementary transposition**.In cyclic notation we denote

$$\tau_{i,i+1} = (i, i+1)$$

Theorem 76

Every permutation can be written as a product

$$\sigma = \tau_1 \circ \tau_2 \ldots \circ \tau_m$$

where each τ_i is an elementary transposition

Theorem 77

Every permuation σ can be wrriten as either a product of an even number of elementary transpositions or as product of an odd number of elementary transpositions but not both

Definition 78 (Sign of Permuation)

Let $x_1, \ldots, x_k$ be the coordinate functions on $\mathbb{R}^k$. For $\sigma \in S_k$ we define

$$(-1)^\sigma = \prod_{i<j} \frac{x_{\sigma(i)} - x_{\sigma(j)}}{x_i - x_j}$$

Proposition 79

Then $(-1)^\sigma = \pm 1$

Proof. Assume all $i < j$. If $p = \sigma(i) < \sigma(j) = q$, then the $x_p - x_q$ here in the numerator happens once because corresponding pairs of $x_i - x_j$ in the denoinator do not repeat. Moreover $x_p - x_q$ will occur once in the denominator too because p, q can be a possible pair of $i < j$. If $p = \sigma(i) > \sigma(j) = q$, like the above $x_p - x_q$ will occur once in the denominator. However it is the negative that is $x_q - x_p$ that will occur once in the denominator since $q > p$ now instead. Therefore in the first case, this corresponds to a contribution of $\frac{x_p - x_q}{x_p - x_q} = 1$ to the product. In the second case, this contributes to a $\frac{x_p - x_q}{x_q - x_p} = -1$ to the product. Therefore the magnitude of the total product is clearly 1 and will be negative if an odd number of the second case occurs.

Proposition 80

If τ is a transposition then $(-1)^\tau = -1$

Proof. Referring to 79, clearly a transposition corresponds to a contribution of $\frac{x_p - x_q}{x_q - x_p} = -1$ to the product which is the second case.

Corollary 81

If σ is the product of an odd number of transpositions then $(-1)^\sigma = -1$ and if σ is an even number of transpositions then $(-1)^\sigma = +1$

Proof. This is clear from the above propositions

Proposition 82

If $\sigma = \tau_1 \circ \ldots \circ \tau_m$ where τ_i are elementary transpositions then the sign of the permutation is given by

$$\operatorname{sgn} \sigma = (-1)^\sigma = (-1)^m$$

Definition 83 (Permuation on tensors)

Define the map

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

where clearly $T^\sigma \in \mathcal{L}^k$ To see why we take inverses consider the example below

Example 84

Consider

$$\sigma(1) = 2 \quad \sigma(2) = 3 \quad \sigma(3) = 1$$

or

$$\sigma = (1, 2, 3)$$

Correct:

$$T(1, 2, 3) \xrightarrow{\sigma} T(3, 1, 2) = T^{\sigma}(1, 2, 3)$$

Wrong:

$$T(1, 2, 3) \xrightarrow{\sigma} T(2, 3, 1) = T^{\sigma}(1, 2, 3)$$

In the wrong approach you replaced 3 with 1 , 2 with 3 , 1 with 2 in the tuple (1,2,3). However it is a shift in a position that should be done instead $(1, 2, 3) \to (3, 2, 1)$ which is actually done using σ^{-1} instead as stated in the definition.

The confusion arose because you know functions are basically maps from one element to another. However the element in question here is the *position* not the *value*. The map tells you to go from position x to position y. Therefore replacing an element at position x with an element from position y is exact opposite of what the function implies because x is the starting point not the end point! To sum up we have

$$\sigma(x) \to y \quad \Leftrightarrow \quad v_y \to v_x \quad \equiv \quad v_y \to v_{\sigma^{-1}(y)}$$

Where the left is the position map while the right is the value map.

Theorem 85

$$(T^{\sigma})^{\tau} = T^{\tau\sigma}$$

Proof.

$$\begin{aligned} T^{\tau\sigma}(v_1, \ldots, v_k) &= T(v_{\sigma^{-1}\tau^{-1}(1)}, \ldots, v_{\sigma^{-1}\tau^{-1}(k)}) \\ &= T^{\sigma}(v_{\tau^{-1}(1)}, \ldots, v_{\tau^{-1}(k)}) \\ &= (T^{\sigma})^{\tau}(v_1, \ldots, v_k) \end{aligned}$$

Fact 86

Suppose $T \in \mathcal{L}^k$ is decomposable

$$T^{\sigma}(v_1, \ldots, v_k) = T(v_{\sigma^{-1}(1)}, \ldots v_{\sigma^{-1}(k)}) = \ell_1(v_{\sigma^{-1}(1)}) \ldots \ell_k(v_{\sigma^{-1}(k)}) = \ell_{\sigma(1)}(v_1) \ldots \ell_{\sigma(k)}(v_k) = \ell_{\sigma(1)} \otimes \ldots \otimes \ell_{\sigma(k)}$$

Because notice that $\sigma^{-1}(i) = j$ and $i = \sigma(j)$, writing in ascending order of j so that v_j corresponds to correct positional arguments in T^{σ} we arrive at the final equality which is the decomposition of T^{τ}. Note I am not implying that the products of ℓ here are commutative, recall that tensors products are not commutative! There is only an equality in terms of the output in $\mathbb{R}$ since $\mathcal{L}^k : V^k \to \mathbb{R}$.

Definition 87

A tensor $T \in \mathcal{L}^k$ is alternating if

$$T^\sigma = (-1)^\sigma T$$

for all $\sigma \in S_k$ The set of all alternating k-tensors is denoted by $\mathcal{A}^k(V)$

Definition 88

Given a k-tensor $T \in \mathcal{L}^k(V)$ we define the alternating operator $\text{Alt} : \mathcal{L}^k(V) \to \mathcal{A}^k(V)$ by

$$\text{Alt}(T) = \sum_{\tau \in S_k} (-1)^\tau T^\tau$$

sum over $\tau \in S_k$ clearly means sum over all $k!$ permutations of τ

Theorem 89

The alternating operator has the following properties

(a) $\text{Alt}(T) \in \mathcal{A}^k(V)$

(b) $T \in \mathcal{A}^k(V)$ then $\text{Alt}(T) = k!T$

(c) $(\text{Alt}(T))^\sigma = \text{Alt}(T^\sigma) = (-1)^\sigma \text{Alt}(T)$

(d) The map $\text{Alt} : \mathcal{L}^k(V) \to \mathcal{A}^k(V)$ is linear

Proof. For (a) since $\text{Alt}(T) = \sum_{\tau \in S_k} (-1)^\tau T^\tau$ so

$$\begin{aligned}
\text{Alt}(T)^\sigma = (\text{Alt}(T))^\sigma = \left(\sum_\tau (-1)^\tau (T^\tau) \right)^\sigma &= \sum_\tau (-1)^\tau (T^\tau)^\sigma && (1)\\
&= \sum_\tau (-1)^\tau (T^{\sigma\tau}) && (2)\\
&= (-1)^\sigma \sum_\tau (-1)^\sigma (-1)^\tau (T^{\sigma\tau}) && (3)\\
&= (-1)^\sigma \sum_{\sigma\tau} (-1)^{\sigma\tau} (T^{\sigma\tau}) && (4)\\
&= (-1)^\sigma \text{Alt}(T) && (5)
\end{aligned}$$

(1) follows because σ permutes arguments of T.The signum function is only dependent on τ, it is not affected by changes in order of arguments of T.(3) follows because $(-1)^\sigma(-1)^\sigma = 1$.(4) follows because a sum over $\tau \in S_k$ also sums possible $\sigma \circ \tau \in S_k$, holding σ fixed. Therefore by definition of Alt (5) follows. For (b), consider $\text{Alt}(T) = \sum_{\tau \in S_k} (-1)^\tau T^\tau$ but $T^\tau = (-1)^\tau T$ so combining we have

$$\text{Alt}(T) = \sum_\tau (-1)^\tau (-1)^\tau T = \sum_\tau T = k!T$$

For (c) consider

$$\begin{aligned}
\operatorname{Alt}(T^\sigma) &= \sum_\tau (-1)^\tau (T^\sigma)^\tau && (1)\\
&= \sum_\tau (-1)^\tau (T^{\tau\sigma}) && (2)\\
&= (-1)^\sigma \sum_\tau (-1)^\sigma(-1)^\tau (T^{\tau\sigma}) && (3)\\
&= (-1)^\sigma \sum_{\tau\sigma} (-1)^{\tau\sigma} (T^{\tau\sigma}) && (4)\\
&= (-1)^\sigma \operatorname{Alt}(T) = (\operatorname{Alt}(T))^\sigma && (5)
\end{aligned}$$

where (5) follows from our results in (a).

Definition 90

An *individual* multi index $I = (i_1, \ldots, i_k)$ is **repeating** if $i_r = i_s$ and $r \neq s$. A multi index I is **strictly increasing** if all its elements $1 \leq i_1 < \ldots i_k \leq n$.(note the strict inequality between elements). Given $\sigma \in S_k$ we denote

$$I^\sigma = (i_{\sigma(1)} \ldots i_{\sigma(k)})$$

Theorem 91

If J is a non-repeating multi-index then there exists a permutation σ such that $J = I^\sigma$ where I is strictly increasing. Therefore we define

$$e_J^* = e_{I^\sigma}^* = e_{\sigma(i_1)}^* \otimes \ldots \otimes e_{\sigma(i_k)}^* = (e_I^*)^\sigma$$

Define $\psi_I = \operatorname{Alt}(e_I^*) = \sum_\sigma (-1)^\sigma (e_I^*)^\sigma$. Then we have the following properties

1. $\psi_{I^\sigma} = (-1)^\sigma \psi_I$
2. If I is repeating then $\psi_I = 0$
3. If I, J are strictly increasing then

$$\psi_I(e_{j_1}, \ldots, e_{j_k}) = \begin{cases} 1 & I = J \\ 0 & I \neq J \end{cases}$$

Proof. For (1) consider

$$\begin{aligned}
\psi_{I^\sigma} &= \operatorname{Alt}(e_{I^\sigma}^*)\\
&= \operatorname{Alt}((e_I^*)^\sigma)\\
&= (-1)^\sigma \operatorname{Alt} e_I^*\\
&= (-1)^\sigma \psi_I
\end{aligned}$$

For (2) Suppose that I is repeating then there obviously exists $I = T^\tau$ for some transposition τ. But recall from 79 that $(-1)^\tau = -1$. Therefore we have

$$\psi_I = \psi_{I^\tau} = (-1)^\tau \psi_I = -\psi_I$$

So we have $\psi_I = 0$. For (3) consider

$$\psi_I = \operatorname{Alt}(e_I^*) = \sum_{\tau} (-1)^{\tau} \underbrace{e_{I^{\tau}}^*(e_{j_1}, \dots, e_{j_k})}_{\begin{cases} 1 & I^{\tau} = J \\ 0 & I^{\tau} \neq J \end{cases}}$$

However $I^{\tau} = J$ if and only τ is the identity permutation because I and J are both strictly increasing so we can replace I^{τ} above with I.

Corollary 92

The alternating k-tensors ψ_I where I is strictly increasing are a basis of $\mathcal{A}^k$

Proof. Recall from 70 that any k-tensor T can be expanded as $T = \sum c_I e_I^*$. Recall from 70 that $\operatorname{Alt}(T) \in \mathcal{A}^k$ therefore any alternating k-tensor can be expressed as

$$\operatorname{Alt}(T) = \sum c_I \operatorname{Alt}(e_I^*) = \sum_I c_I \psi_I$$

since $\operatorname{Alt}(T)$ is linear. From 70 again we know that if I is repeating then $\psi_I = 0$. But if I is non-repeating and not strictly increasing there exists $I = J^{\sigma}$ where J is strictly increasing so we have $\psi_I = \psi_{J^{\sigma}} = (-1)^{\sigma} \psi_J$. Therefore we can just consider only strictly increasing and still span all cases as desired. Now we have to show linear indepedence. Let $T \in \mathcal{A}^k$ be the zero function Then for any strictly increasing J we must have $T(e_{j_1}, \dots, T(e_{j_k})) = c_j$ therefore the only solution to this is the trivial sum as desired.

Theorem 93

If $0 \leq k \leq n$ then

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

and if $k > n$ then $\mathcal{A}^k = \{0\}$

Proof. Think of the sum of strictly increasing I in the above derivation. Now consider $\frac{n!}{(n-k)!}$ counts all possible non-repeating I. Then dividing by $k!$ is to ignore all permutations of each non-repeating I found. Because we only want one permutation from each I, that is the one that is strictly increasing.

Definition 94 (redundant k-tensors)

Let $T \in \mathcal{L}^k$ be a decomposable vector $T = \ell_1 \otimes \dots \otimes \ell_k$ then the k-tensor T is **redundant** if $\ell_i = \ell_{i+1}$ for some $1 \leq i \leq k-1$ we denote

$$\mathcal{I}^k \equiv \{\text{redundant k-tensors}\}$$

Theorem 95

If $T \in \mathcal{I}^k$ then $\operatorname{Alt}(T) = 0$

Proof. if $\ell_i = \ell_{i+1}$ then there exists $T^{\tau} = T$ but

$$\operatorname{Alt}(T^{\tau}) = (-1)^{\tau} \operatorname{Alt}(T) = -\operatorname{Alt}(T)$$

Theorem 96

Suppose $T \in \mathcal{I}^k$ and $T' \in \mathcal{L}^m$ then

$$T' \otimes T \in \mathcal{I}^{k+n}$$

and

$$T \otimes T' \in \mathcal{I}^{k+n}$$

Proof. THe proof is obvious when you assume both are decomposable vectors that take their total tensor product in $T \otimes T'$ and vice versa.

Theorem 97

For each $T \in \mathcal{L}^k$ and $\sigma \in S_k$ there exists some $W \in \mathcal{I}^k$ such that

$$T = (-1)^\sigma T^\sigma + W$$

Proof. By induction let us consider the base case $k = 2$. Then have $T = \ell_1 \otimes \ell_2$ Therefore the only way to create a redundant tensor is to the transposition permutation $\sigma = \tau_{1,2}$. We know that $(-1)^\sigma = -1$. Therefore we can define W by

$$\begin{aligned} W = T + T^\sigma &= (\ell_1 \otimes \ell_2) + (\ell_2 \otimes \ell_1) \\ &= (\ell_1 + \ell_2)(\ell_1 + \ell_2) - \ell_1 \otimes \ell_1 - \ell_2 \otimes \ell_2 \in \mathcal{I}^2 \end{aligned}$$

Since notice this is a sum of tensor products of identical ℓ. Suppose we now have the transposition permutation $\sigma = \tau_{i,i+1}$ we can express for any k by the associativity and distributive properties of tensor products(which are not commutative)

$$\begin{aligned} W = T + T^\sigma &= (\ell_1 \ldots)(\ell_i \otimes \ell_{i+1})(\ldots \ell_k) + (\ell_1 \ldots)(\ell_{i+1} \otimes \ell_i)(\ldots \ell_k) \\ &= (\ell_1 \ldots)(\ell_i \otimes \ell_{i+1} + \ell_{i+1} \otimes \ell_i)(\ldots \ell_k) \end{aligned}$$

To prove the induction step suppose $\beta = \tau_2 \ldots \tau_r$ and $\sigma = \tau_1 \beta$

$$T^\sigma = (T^\beta)^{\tau_1}$$

which by induction hypoethesis is true so we have

$$T^\beta = (-1)^\beta T + W$$

Corollary 98

For every $T \in \mathcal{L}^k$

$$\mathrm{Alt}(T) = k! T + W$$

for some $W \in \mathcal{I}^k$

Proof. Simply substitute the above theorem into Alt

$$\text{Alt}(T) = \sum_{\sigma}(-1)^{\sigma}T^{\sigma} = \sum_{\sigma}(-1)^{\sigma}((-1)^{\sigma}(T - W))$$

since $\frac{1}{(-1)^{\sigma}} = (-1)^{\sigma}$ and the conclusion follows because $\sum_{\sigma} -W \in \mathcal{I}^k$ clearly

Theorem 99

Every $T \in \mathcal{L}^k$ can be written uniquely as the sum

$$T = T_1 + T_2$$

where $T_1 \in \mathcal{A}^k$ and $T_2 \in \mathcal{I}^k$

Proof. From the above we already know that $\text{Alt}(T) = k!T + W$ for some $W \in \mathcal{I}^k$. Then therefore we have

$$T = \frac{1}{k!}\text{Alt}(T) - \frac{1}{k!}W = T_1 + T_2$$

to show uniqueness, we have to prove that $k!T_1 = \text{Alt}(T)$ and $W = -T_2$. To do so take Alt on all sides, obtaining

$$\text{Alt}(T) = \text{Alt}(T_1) + \text{Alt}(T_2)$$

but since $T_2 \in \mathcal{I}^k$ hence $\text{Alt}(T_2) = 0$. Since $T_1 \in \mathcal{A}^k$ we have $\text{Alt}(T_1) = k!T_1$. Therefore we have shown $k!T_1 = \text{Alt}(T)$ which also implies T_2 is unique.

Corollary 100

$\mathcal{I}^k = \ker \text{Alt}$ and $\mathcal{A}^k \cap \mathcal{I}^k = \{0\}$

Proof. If $\text{Alt}\, T = 0$ then

$$T = -\frac{1}{k!}W, W \in \mathcal{I}^k$$

so $T \in \mathcal{I}^k$. The second can be proven by contradiction. If this is false, then there exists $T = T_1 + T_2 \in \mathcal{L}^k$ where $T_1 = T_2 \neq 0 \in \mathcal{A}^k \cap \mathcal{I}^k$. However we know that $k!T_1 = \text{Alt}(2T_1) = 0$ since $T_1 \in \mathcal{I}^k$ but we had assumed $T_1 \neq 0$ so we have a contradiction.

Definition 101

The space $\mathcal{I}^k$ is a subspace of $\mathcal{L}^k$ since it is obviously closed under addition,scalar multiplication and contains indentity zero(the zero tensor). There we can define the **quotient space** by

$$\Lambda^k(V^*) \equiv \mathcal{L}^k(V)/\mathcal{I}^k(V)$$

.

Fact 102
Consider

$$\Lambda^1(V^*) \equiv \mathcal{L}^1/\mathcal{I}^1$$

But the $\mathcal{I}^1 = \{0\}$ since there is clearly no way to permute a k-1 tensor with itself. Therfore

$$\Lambda^1(V^*) \equiv \mathcal{L}^1/\mathcal{I}^1 \equiv \mathcal{L}^1 \equiv V^*$$

Theorem 103
The map defined by $\pi : \mathcal{L}^k \to \mathcal{L}^k/\mathcal{I}^k$ whose nullspace is $\ker \pi = \mathcal{I}^k$. Then $\pi|\mathcal{A}^k$ is bijective that is

$$\mathcal{A}^k(V) \simeq \Lambda^k(V^*)$$

Proof. For every element of Λ^k can write it as some $\pi(T)$ where $T \in \mathcal{L}^k$ then we have by linearity

$$\pi(T) = \pi(T_1) + \pi(T_2) = \pi(T_1)$$

So π maps $\mathcal{A}^k$ onto Λ^k, that is surjective. Since from previously we know that $\mathcal{A}^k \cap \mathcal{I}^k = \{0\}$, that is the nullspace of the map $\pi : \mathcal{A}^k \to \Lambda^k$ only contains the zero vector therefore we have injectivity.

3.3 wedge product

Recall map π maps $\mathcal{A}^k$ bijectively into Λ^k.

Definition 104 (wedge product)
Then we define Let $\mu_i \in \Lambda^{k_i}, i = 1, 2$ so $\mu_i = \pi(T_i)$ for some $T_i \in \mathcal{L}^{k_i}$. Define $k = k_1 + k_2$ so $T_1 \otimes T_2 \in \mathcal{L}^k$. Then we define the **wedge product** by

$$\pi(T_1 \otimes T_2) = \mu_1 \wedge \mu_2 \in \Lambda^k$$

Proposition 105
The wedge product is well defined(does not depend on choice of T) and it is associative as well as distributive. That is to say if $k_1 = k_2$

$$(\mu_1 + \mu_2) \wedge \mu_3 = \mu_1 \wedge \mu_3 + \mu_2 \wedge \mu_3$$

$$\mu_1 \wedge (\mu_2 + \mu_3) = \mu_1 \wedge \mu_2 + \mu_1 \wedge \mu_3$$

$$(\mu_1 \wedge \mu_2) \wedge \mu_3 = \mu_1 \wedge (\mu_2 \wedge \mu_3)$$

Proof. Let $T_1' = T_1 + W_1$ and $T_2' = T_2 + W_2$ where $T_i' \in \mathcal{L}^{k_i}$ too. Then we have

$$T_1' \otimes T_2' = T_1 \otimes T_2 + \underbrace{W_1 \otimes T_2 + T_1 \otimes W_2 + W_1 \otimes W_2}_{\in \mathcal{I}^k}$$

Therefore

$$\mu_1 \wedge \mu_2 = \pi(T_1' \otimes T_2') = \pi(T_1 \otimes T_2)$$

The other properties follow the associative and distributive properties of tensor products then the linearity of π.

Definition 106

The element $\mu \in \Lambda^k$ is **decomposable** if it is of the form

$$\mu = \ell_1 \wedge \ldots \wedge \ell_k$$

where each $\ell_i \in \Lambda^1 = V^*$ That means

$$\mu = \pi(\ell_1 \otimes \ldots \otimes \ell_k)$$

Proposition 107

Take a permutation $\sigma \in S_k$ and $w \in \Lambda^k$ such that $w = \pi(T)$ where $T \in \mathcal{L}^k$

$$w^\sigma = \pi(T^\sigma)$$

Show that is w^σ is the same value regardless of T for the same w found and that $w^\sigma = (-1)^\sigma w$

Proof. To show that w^σ does not depend on the choice of T. Let $w = \pi(T) = \pi(T')$. Then we know $\pi(T - T') = 0$ meaning $T - T' \in \mathcal{I}^k$. Therefore recall we can do for some $W \in \mathcal{I}^k$

$$(T - T')^\sigma = (-1)^\sigma (T - T') + W$$

Therefore we have $(T')^\sigma - T^\sigma \in \mathcal{I}^k$ so we have

$$w^\sigma = \pi((T')^\sigma) = \pi(T^\sigma)$$

For the latter it is cleary when you consider

$$w^\sigma = \pi(T^\sigma) = \pi((-1)^\sigma T + W) = (-1)^\sigma \pi(T) = (-1)^\sigma w$$

Remark 108. *This is necessary because π is only bijective when restricted to domain of $\mathcal{A}^k$ but now we are taking $\mathcal{L}^k$ in general!*

Corollary 109

Suppose w is decomposable where we have $w = \ell_1 \wedge \ldots \wedge \ell_k, \ell_i \in \mathcal{L}^1 = V^*$ Then

$$w = \pi(\ell_1 \otimes \ldots \otimes \ell_k)$$

therefore from the previous proposition we have

$$\begin{aligned} w^\sigma &= \pi((\ell_1 \otimes \ldots \otimes \ell_k)^\sigma) \\ &= \pi(\ell_{\sigma(1)} \otimes \ldots \otimes \ell_{\sigma(k)}) \\ &= \ell_{\sigma(1)} \wedge \ldots \wedge \ell_{\sigma(k)} \\ &= (-1)^\sigma \ell_1 \wedge \ldots \wedge \ell_k \end{aligned}$$

Corollary 110
Consider the following

1. if $\mu \in \Lambda^2$ and $\ell \in \Lambda^1$ then
$$\mu \wedge \ell = \ell \wedge \mu$$
2. if $\mu \in \Lambda^2$ and $\nu \in \Lambda^2$ then
$$\mu \wedge \nu = \nu \wedge \mu$$
3. if $\mu \in \Lambda^k$ and $\nu \in \Lambda^\ell$ then
$$\mu \wedge \nu = (-1)^{k\ell} \nu \wedge \mu$$

Proof. For (1) consider

$$\begin{aligned}(\ell_1 \wedge \ell_2) \wedge \ell_3 &= \ell_1 \wedge (\ell_2 \wedge \ell_3) \\ &= -\ell_1 \wedge (\ell_3 \wedge \ell_3) \\ &= \ell_3 \wedge (\ell_1 \wedge \ell_3)\end{aligned}$$

where the sign in the middle is reversed because 1 transposition happened then another happened in the last so $(-1)(-1) = 1$ For (2) consider

$$(\ell_1 \wedge \ell_2) \wedge (\ell_3 \wedge \ell_4) = (\ell_3 \wedge \ell_4) \wedge (\ell_1 \wedge \ell_2)$$

Consider in the LHS both ℓ_3 and ℓ_4 took 2 swaps each to get to the form in the RHS when you move them in ascending order(that is move ℓ_3 to the correct postition first before ℓ_4.) For (3) the logic follows similary to (2). In the LHS every ℓ_i elements on in ν has to move k positions to the left. And when you move all ℓ in ν in ascending order you eventually get a total of $k\ell$ swaps.

Theorem 111
$\tilde{e}_I = \pi(e_I^*)$ is a basis of $\Lambda^k(V^*)$

Proof. We already know that $\mathcal{A}^k$ maps bijectively with Λ^k, we already know the number of basis functions are the same. We know that any element in Λ^k can be expressed as $\pi(\sum c_I \psi_I)$. So all we need to do just to find the 1 to 1 equivalent form of the basis functions in Λ^k. That is we just need to analyze one basis function of $\mathcal{A}^k$

$$\begin{aligned}\pi(\psi_I) &= \sum_{\sigma \in S_k} (-1)^\sigma \pi((e_I^*)^\sigma) \\ &= \sum (-1)^\sigma (-1)^\sigma \pi(e_I^*) \\ &= k! \pi(e_I^*)\end{aligned}$$

3.4 dual transformation $\Lambda^k(W^*) \to \Lambda^k(V^*)$

Proposition 112 (Dual Transformation of alternating tensors)
Let $A^* : \mathcal{L}^k(W) \to \mathcal{L}^k(V)$ and $A : V \to W$ Let $\mu \in \Lambda^k(W^*)$ so $\mu = \pi(T)$ for some $T \in \mathcal{L}^k(W)$.Then we apply pullback operators A^* like so

$$A^*\mu = \pi(A^*T) \in \Lambda^k(V^*)$$

Prove that this does not depend on the choice of T.

Proof. Like before we begin with $\mu = \pi(T) = \pi(T')$ so $T - T' \in \mathcal{I}^k(W)$. Then applying pullback operators which are linear as proven earlier we have

$$A^*T' - A^*T \in \mathcal{I}^k(V)$$

which shows that $A^*\mu = \pi(A^*T') = \pi(A^*T)$

Proposition 113
Pullback operators on wedge products are still linear and distributive on wedge products

$$A^*(\mu_1 \wedge \mu_2) = A^*\mu_1 \wedge A^*\mu_2$$

Proof. Consider

$$\pi((F^*\omega) \otimes (F^*\eta)(v_{\sigma(1)}, \dots, v_{\sigma(p+q)})) = \pi((F^*\omega)(v_{\sigma(1)}, \dots, v_{\sigma(p)}) \cdot (F^*\eta)(v_{\sigma(p+1)}, \dots, v_{\sigma(p+q)}))$$

$$= \pi(\omega(F(v_{\sigma(1)}), \dots, F(v_{\sigma(p)})) \cdot \eta(F(v_{\sigma(p+1)}), \dots, F(v_{\sigma(p+q)})))$$

$$= \pi((\omega \otimes \eta)(F(v_{\sigma(1)}), \dots, F(v_{\sigma(p+q)}))) = \pi(F^*(\omega \otimes \eta)((v_{\sigma(1)}), \dots, (v_{\sigma(p+q)})))$$

3.5 determinant

Definition 114
Let V be an n-dimensional vector space.
We call the $\Lambda^n(V^*)$ the **nth exterior power** of V where $\dim \Lambda^n(V^*) = 1$ since $\mathcal{A}^n(V) \simeq \Lambda^n(V^*)$ and that $\binom{n}{n} = 1$.
Suppose the pullback operator was defined by

$$A^* : \Lambda^n(V^*) \to \Lambda^n(V^*)$$

where $A : V \to V$. That is for all $w \in \Lambda^n(V^*)$ we have

$$A^*w = \lambda_A w$$

since we have 1-dimensional invariant map. Then the determinant of A is denoted by

$$\det(A) = \lambda_A$$

Theorem 115

The determinant has the following properties

1. If $A = I$ then $\det(A) = 1$
2. If A, B are linear maps of V into V then $\det(AB) = \det(A)\det(B)$

Proof. Let $w \in \Lambda^n(V^*)$ Then recall 72 we have

$$\begin{aligned}(AB)^* w &= \det(AB) w \\ &= B^*(A^* w) \\ &= B^*(\det A) w \\ &= \det(A)\det(B) w\end{aligned}$$

Problem 116

Let W, V be n-dimensional vector spaces and let $A : V \to W$ be a linera map. We have the bases

- $e_1, \ldots e_n$ basis of V
- $e_1^*, \ldots e_n^*$ dual basis of $V^* = \mathcal{L}(V)$
- $f_1, \ldots f_n$ basis of W
- $f_1^*, \ldots f_n^*$ dual basis of $W^* = \mathcal{L}(W)$

Recall from linear algebra we can write $AX = Y$ so we have

$$Ae_i = \sum a_{ij} f_j$$

where the RHS is exactly how y is calculated. So that matrix of the transformatino is $A \sim [a_{ij}]$. Now find a relation for the matrix of transformation for

$$A^* w = \det[a_{ij}] v$$

that is $\det(A) = \det(a_{ij})$

Proof. First express the pullback operation

$$A^* f_j^* = \sum a_{jk} e_k^*$$

and take $w = f_1^* \wedge \ldots \wedge f_n^* \in \Lambda^n(W^*)$ and $v = e_1^* \wedge \ldots \wedge e_n^* \in \Lambda^n(V^*)$ which we know is basis vector of $\Lambda^n(W^*)$ and

$\Lambda^n(V^*)$ respectively. Therefore

$$
\begin{aligned}
A^*(f_1^* \wedge \ldots \wedge f_n^*) &= A^* f_1^* \wedge \ldots \wedge A^* f_n^* && (6)\\
&= \left(\sum_{k_1}^{n} a_{1,k_1} e_{k_1}^*\right) \wedge \ldots \wedge \left(\sum_{k_n}^{n} a_{n,k_n} e_{k_n}^*\right) && (7)\\
&= \sum_{k_1,\ldots,k_n} (a_{1,k_1} \ldots a_{n,k_n})(e_{k_1}^* \wedge \ldots \wedge e_{k_n}^*) && (8)\\
&= \sum_{\sigma} (a_{1,\sigma(1)} \ldots a_{n,\sigma(n)})(e_{\sigma(1)}^* \wedge \ldots \wedge e_{\sigma(n)}^*) && (9)\\
&= \sum_{\sigma} (-1)^{\sigma} (a_{1,\sigma(1)} \ldots a_{n,\sigma(n)})(e_1^* \wedge \ldots \wedge e_n^*) && (10)
\end{aligned}
$$

where (10) follows from 109 and (7) follows from linearity of pullback operators as proven earlier.(9) follows because all repeating $\pi(e_{k_1}^* \otimes \ldots \otimes e_{k_n}^*)$ if they are redundant tensors then $\pi = 0$ so we only consider non-repeating multidexes which is precisely the permutations!

$$\det A = \sum_{\sigma} (-1)^{\sigma} (a_{1,\sigma(1)} \ldots a_{n,\sigma(n)})$$

since that is precisely the definition of determinant of a square matrix

Fact 117

Let A be an $n \times n$ matrix then

$$\det A = \sum_{\sigma} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \ldots a_{n,\sigma(n)}$$

3.6 tangent spaces and k-forms

Definition 118

The **tangent space** of $p \in \mathbb{R}^n$ is

$$T_p\mathbb{R}^n = \{(p, v) : v \in \mathbb{R}^n\}$$

where

$$T_p\mathbb{R}^n \simeq \mathbb{R}^n$$

and $(p, v) \to v$ where $p, v \in \mathbb{R}^n$

Fact 119

Let U be an open set in $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^m$ be a $\mathcal{C}^1$ map. Also let $p \in U$ and define $q = f(p)$. We define a linear map

$$df_p : T_p\mathbb{R}^n \to T_q\mathbb{R}^m$$

which is related by the following commutative diagram

$$\begin{tikzcd} T_p\mathbb{R}^n \arrow[r, "df_p"] \arrow[d, "\simeq"'] & T_q\mathbb{R}^m \\ \mathbb{R}^n \arrow[r, "Df(p)"] & \mathbb{R}^m \arrow[u, "\simeq"'] \end{tikzcd}$$

therefore note the following notation

$$df_p(p, v) = (q, Df(p)v) = \left(\sum_i \frac{\partial f}{\partial x_i}(p)(dx_i)_p\right)(v)$$

$$df_p(p, e_i) = (q, Df(p)e_i) = \frac{\partial f}{\partial x_i}(p)$$

$$(p, e_i) = \left(\frac{\partial}{\partial x_i}\right)_p$$

Note that the corresponding 1 form $df : p \to df_p$ is written with 2 arguments one for (x) and one for the tuple pair (x, v)

$$df(x)(x, v) = df_x(x, v) = (y, Df(x)v) = \left(\sum_i \frac{\partial f}{\partial x_i}(x)(dx_i)_x\right)(v)$$

$$df_p(p, e_i) = (q, Df(p)e_i) = \frac{\partial f}{\partial x_i}(p)$$

where $y = f(x)$. Note that is seems that sometimes we just omit the q in $(q, Df(p)v)$ in calculations and just write

$$Df(p)v = \left(\sum_i \frac{\partial f}{\partial x_i}(p)(dx_i)_p\right)(v)$$

because it is quite obvious where the root is. $f(p) = q$!!

Definition 120

The **cotangent space** of $\mathbb{R}^n$ at p is the space

$$T_p^*\mathbb{R}^n \equiv (T_p\mathbb{R}^n)^*$$

which is the dual of the tangent space of $\mathbb{R}^n$ at p that is

$$w : T_p\mathbb{R}^n \to \mathbb{R}$$

where $w \in T_p^*\mathbb{R}^n$

Definition 121

Let U be an open subset of $\mathbb{R}^n$. A **k-form** on U is a function w which assigns to every point $p \in U$ an element w_p of $\Lambda^k(T_p^*\mathbb{R}^n)$(the kth exterior power of $T_p^*\mathbb{R}^n$. That is

$$w : p \to w_p$$

Definition 122

The zero form on an open set $A \in \mathbb{R}^n$ is the function $f : A \to \mathbb{R}$ be a function of $\mathcal{C}^r$. The 1-form denoted by df on A is *defined* by

$$df(x)(x; v) = f'(x; v) = Df(x) \cdot v$$

Let x be a point in $\mathbb{R}^n$ and $x_i(x) = xe_i$ be its coordinate function then

$$dx_i(x)(x; v) = Dx_i(x) \cdot v$$

Then

$$dx_I = dx_{i_1} \wedge \ldots \wedge dx_{i_k}$$

Let I be strictly increasing multi-indexes. Then the k form is defined by

$$w = \sum_I b_I dx_I$$

and the forms dx_I and dx_i are defined by

$$dx_i(x)(x; v) = Dx_i(x) \cdot v = [0 \ldots 1 \ldots 0] \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = v_i$$

where $\cdot$ indicates the matrix product like before and that since

$$dx_i(x) = Dx_i(x) = \underbrace{\left[De_j x_i(x) = D_j x_i(x) = \frac{\partial x_i}{\partial x_j} \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \right]}_{\text{every j is one column}}$$

Also

$$\begin{aligned} dx_I(x)((x, \mathbf{v}_1) \ldots, (x, \mathbf{v}_k)) &= dx_1(x)(x, \mathbf{v}_1) \wedge \ldots \wedge dx_k(x)(x, \mathbf{v}_k) \\ &= v_{1,1} \ldots v_{k,k} \\ &= \det X_I \end{aligned}$$

Definition 123

As usual we assume the $\cdot$ here to be the matrix product. The wedge product of 2 zero forms f an g are related by

$$f \wedge g = f \cdot g$$

which is actually just the usual product o real valued functions. and the wedge product of 0-form f and the k-form w is related by

$$(w \wedge f)(x) = (f \wedge w)(x) = f(x) \cdot w(x)$$

which is the usual product of the tensor $w(x)$ and the scalar $f(x)$.
However unlike typical wedge products, commutativity holds because scalar fields are forms of order zero that is

$$f \wedge g = (-1)^0 g \wedge f \quad \text{and} \quad f \wedge w = (-1)^0 w \wedge f$$

Example 124

Let $f, g \in \mathcal{C}^\infty = \Omega^0$. Then $w = df \wedge dg$ is the *2-form* that maps

$$p \in U \to df_p \wedge dg_p = w_p \in \Lambda^2(T_p^*\mathbb{R}^n)$$

where each$(df_i)_p \in T_p^*\mathbb{R}^n$. Hmm...I am beginning to see the motivation for the notation of *quotient spaces*

Example 125

Let $e_1, \ldots, e_n$ be the standard basis oof $\mathbb{R}^n$. Let $p \in U$ and let $v_i = (p, e_i) = \left(\frac{\partial}{\partial x_i}\right)_p$. Then $v_1, \ldots, v_n$ are a basis of $T_p\mathbb{R}^n$ since $\mathbb{R}^n \simeq T_p\mathbb{R}^n$. We now attempt to find the basis for its dual, $T_p^*\mathbb{R}^n$. Everything below is analagous to 61

$$\lambda_i = df_p(v_i) = df_p(p, e_i) = Df(p)e_i = \frac{\partial f}{\partial x_i}(p)$$

where x_i is the ith coordinate function.That is $pe_i = x_i$ Like previously where we had $e_i^*(e_i)$ we define

$$(dx_i)_p(v_j) = (dx_i)_p(p, e_j) = Dx_i(p)e_j = \frac{\partial x_i}{\partial x_j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

so we can wite every df_p and $v = a_1 v_1 + \ldots + a_n v_n$

$$df_p(v) = \sum_i \lambda_i = \left(\sum_i \lambda_i (dx_i)_p\right)(v_i) = \left(\sum_i \frac{\partial f}{\partial x_i}(p)(dx_i)_p\right)(v)$$

where $(dx_1)_p, \ldots, (dx_n)_p$ is the basis for $T_p^*\mathbb{R}^n$. So we can also write

$$df = \left(\sum_i \frac{\partial f}{\partial x_i}(dx_i)\right)$$

Also the basis of $\Lambda^k(T_p^*\mathbb{R}^n)$ is

$$(dx_I)_p = (dx_{i_1})_p \wedge \ldots (dx_{i_k})_p$$

where recall from earlier I must be a strictly increasing multi-index Therefore finally in summary we have

$$w_p = \sum a_I(p)(dx_I)_p, \quad a_I(p) \in \mathbb{R}$$

and

$$w = \sum a_I(dx_I)$$

where I is strictly increasing.The k-form w is $\mathcal{C}^r(U)$ if each $a_I \in \mathcal{C}^r(U)$

Definition 126

We define

$$\Omega^k(U) = \text{the set of all } \mathcal{C}^\infty \text{ k-forms}$$

If $w \in \Omega^k(U)$ then w can be written uniquely as

$$w = \sum f_I dx_I = \sum_I f_I \wedge dx_I$$

where I is strictly increasing then we *define*

$$dw = \sum df_I \wedge dx_I = \sum_I \left(\sum_j^n (D_j f) dx_j \right) \wedge dx_I \in \Omega^{k+1}(U)$$

$$d(dw) = \sum d(df_I) \wedge dx_I = \sum_I \left(\sum_i \left(\sum_j^n D_i D_j f \, dx_i \right) \wedge dx_j \right) \wedge dx_I \in \Omega^{k+2}(U)$$

which makes sense since $f_I \in C^\infty = \Omega^0(U)$. This motivates our definition below. Moreover you will soon know that in fact $d(dw)$ and higher orders of differential operators $d(d \ldots d(dw))$ are all equal zero. This is why there is no chain rule since

$$dw = df_I \wedge dx_I + \underbrace{f_I \wedge d(dx_I)}_{=0}$$

Definition 127

Let U be an open set in $\mathbb{R}^n$. For each $k = 0, \ldots, n-1$ we define the differential operator

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

Specifically for the zero-form $f : U \in \mathbb{R}^n \to \mathbb{R}$ we have

$$d : \Omega^0(U) \to \Omega^1(U) \quad \Leftrightarrow \quad f \to df$$

and that $\mathcal{C}^\infty(U) = \Omega^0(U)$. Then we have

$$df_p = T_p\mathbb{R}^n \to T_c\mathbb{R} = \mathbb{R}$$

where $c = f(p)$. Also clearly $df_p \in T_p^*\mathbb{R}^n = \Lambda^1(T_p^*\mathbb{R}^n)$ and $df \in \Omega^1(U)$
Notice how this contrasts with other functions where we have

$$f : U \in \mathbb{R}^n \to V \in \mathbb{R}^m, f(p) = q \quad \leftrightarrow \quad df_p = T_p\mathbb{R}^n \to T_q\mathbb{R}^m$$

We summarize properties we have for differential operators on $\mathcal{C}^\infty$ below

Theorem 128 (Differential operators on $\mathcal{C}^\infty$)

Consider the following properties

1. Let $f, g \in \mathcal{C}^\infty(U)$ then $d(fg) = gdf + fdg$
2. Also if $a_I \in \mathcal{C}^\infty$ then
$$d(a_I dx_I) = da_I \wedge dx_I$$
3. Finally
$$df = \left(\sum_i \frac{\partial f}{\partial x_i}(dx_i)\right)$$

Proof. For the first property consider

$$D(gh)(x) = gDh(x) + hDg(x)$$

by product rule of differentials. Now multiply all sides by matrix $\cdot$ product v then we get

$$d(gh) = gdh + hdg$$

as desired For the second property, it holds by definition 126 For the third property we have by 122

$$\begin{aligned} df(x)(x; v) = Df(x) \cdot v &= \sum_{i,j} D_j f_i(x) v_j \\ &= \sum_j D_j f(x) v_j \\ &= \sum_j D_j f(x) dx_j(x)(x; v) \end{aligned}$$

Theorem 129

The differential d is linear on 0-forms

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

where f, g are 0-forms

Proof. Let $h = af + bg$ where f, g are 0-forms. Then

$$Dh(x) = aDf(x) + bDg(x)$$

multiplying by matrix product $\cdot$ of v on all sides we have

$$dh(x)(x; v) = adf(x)(x; v) + bdg(x)(x; v)$$

thus

$$dh = adf + bdg$$

□

In fact we can generalize this to k-form.

Theorem 130
The differential d is linear on same k-forms

$$d(w_1 + w_2) = dw_1 + dw_2$$

where w_1, w_2 are k-forms

Proof. Let $w_1 = f_1 dx_1 \wedge \ldots \wedge dx_n$ and $w_2 = f_2 dx_1 \wedge \ldots \wedge dx_n$. Then

$$d(w_1 + w_2) = \sum_j \frac{\partial (f_1 + f_2)}{\partial x_j} dx_j \wedge dx_1 \wedge \ldots \wedge dx_n$$

where you can clearly split into $dw_1 + dw_2$ by linearity of the partial differential

Theorem 131
If $\mu \in \Omega^k(U)$ and $\nu \in \Omega^\ell(U)$ then

$$d(\mu \wedge \nu) = d\mu \wedge \nu + (-1)^k \mu \wedge d\nu$$

Proof. Take $\mu = \sum a_I dx_I$ and $\nu = \sum b_J dx_J$ where I, J are strictly increasing then

$$\mu \wedge \nu = \sum a_I b_J \underbrace{dx_I \wedge dx_J}_{\text{no longer increasing}}$$

Then we have

$$\begin{aligned}
d(\mu \wedge \nu) &= \sum_{i,I,J} \frac{\partial a_I b_J}{\partial x_i} dx_i \wedge dx_I \wedge dx_J && (1)\\
&= \sum_{i,I,J} \frac{\partial a_I}{\partial x_i} b_J dx_i \wedge dx_I \wedge dx_J && (2)\\
&+ \sum_{i,I,J} \frac{\partial b_J}{\partial x_i} a_I dx_i \wedge dx_I \wedge dx_J && (3)
\end{aligned}$$

We were able to split (1) into (2) and (3) due to product rule of differentials. For (2) we have

$$\begin{aligned}
\sum_{i,I,J} \frac{\partial a_I}{\partial x_i} b_J dx_i \wedge dx_I \wedge dx_J &= \left(\sum_{i,I} \frac{\partial a_I}{\partial x_i} dx_i \wedge dx_I \right) \wedge \sum_J b_J dx_J \\
&= d\mu \wedge \nu
\end{aligned}$$

Where we can move the sums around by linearity of tensor product. For (3) we have

$$\begin{aligned}
\sum_{i,I,J} \frac{\partial b_J}{\partial x_i} a_I dx_i \wedge dx_I \wedge dx_J &= (-1)^k \sum_{i,I,J} a_I dx_I \wedge \frac{\partial b_j}{\partial x_i} dx_i \wedge dx_J && (4)\\
&= (-1)^k \left(\sum_I a_I dx_I \right) \wedge \sum_{i,J} \frac{\partial b_J}{\partial_i} dx_i \wedge dx_J && (5)\\
&= (-)^k \mu \wedge d\nu && (6)
\end{aligned}$$

where (5) follows because there are k swaps involved due to the fact that $dx_I = dx_{i_1} \wedge \ldots \wedge dx_{i_k}$. Therefore the combination of our results give the desired conclusion $\square$

Theorem 132
For all $w \in \Omega^k(U)$ we have

$$d(dw) = 0$$

Proof. Let $w = \sum a_I dx_I$ then by definition we have

$$dw = \sum_{j,I} \frac{\partial a_I}{\partial x_j} dx_j \wedge dx_I$$

then

$$d(dw) = \sum_{i,j,I} \frac{\partial^2 a_I}{\partial x_i \partial x_j} dx_i \wedge dx_j j \wedge dx_I$$

we may split this into by the definition above

$$\begin{aligned} d(dw) &= \sum_{i,j,I} \frac{\partial^2 a_I}{\partial x_i \partial x_j} dx_i \wedge dx_j j \wedge dx_I && (1) \\ &= \sum_{i<j,I} \frac{\partial^2 a_I}{\partial x_i \partial x_j} dx_i \wedge dx_j j \wedge dx_I && (2) \\ &+ \sum_{i>j,I} \frac{\partial^2 a_I}{\partial x_i \partial x_j} dx_i \wedge dx_j j \wedge dx_I && (3) \\ &+ \sum_{i=j,I} \frac{\partial^2 a_I}{\partial x_i \partial x_j} dx_i \wedge dx_j j \wedge dx_I && (4) \end{aligned}$$

However we know $dx_i \wedge dx_j = -dx_j \wedge dx_i$ and that the repeating case (4) will be zero since

$$\pi(\ell_i \otimes \ell_j) = \ell x_i \wedge \ell x_j = 0$$

if $i = j$ therefore we have

$$d(dw) = \sum_{i<j,I} \left(\frac{\partial^2 a_I}{\partial x_i \partial x_j} - \frac{\partial^2 a_I}{\partial x_i \partial x_j} \right) dx_i \wedge dx_j j \wedge dx_I = 0$$

Definition 133
A k-form $w \in \Omega^k(U)$ is **decomposable** if

$$w = \mu_1 \wedge \ldots \wedge \mu_k$$

where each $\mu_i \in \Omega^1(U)$

Theorem 134
If w is decomposable then

$$dw = \sum_{i=1}^{k} (-1)^{i-1} \mu_1 \wedge \ldots \wedge \mu_{i-1} \wedge d\mu_i \wedge \mu_{i+1} \wedge \ldots \wedge \mu_k$$

3.7 dual transformation $\Omega^k(W) \to \Omega^k(V)$

Definition 135 (Dual Transformation of forms)
Let U be open in $\mathbb{R}^n$ and V be open in $\mathbb{R}^m$ and $f : U \to V$ be a $\mathcal{C}^\infty$ map
Let $q = f(p), p \in U$
Define the pullback of the map $df_p : T_p\mathbb{R}^n \to T_q\mathbb{R}^m$ by

$$(df_p)^* : \Lambda^k(T_q^*\mathbb{R}^n) \to \Lambda^k(T_p^*\mathbb{R}^m)$$

Note the pullback the positions of p, q are swapped. This is a consequence of chain rule see the example below. That is suppose you have a k-form w on V then as usual we have

$$w \in \Omega^k(V) \to f^*w \in \Omega^k(U)$$

and

$$w_q \in \Lambda^k(T_q^*\mathbb{R}^n) \to f^*w_p \in \Lambda^k(T_p^*)$$

Then upon pullback we have

$$(f^*w)_p = (df_p)^*(w_q) \in \Lambda^k(T_p^*\mathbb{R}^m)$$

Where f^*w is the k-form whose value at $p \in U$ is $(df_p)^*(w_q) = (f^*w)_p$ Specifically we define

$$f^*w(x)((x; v_1), \ldots, (x; v_k)) = w(f(x))(f_*(x; v_1), \ldots, f_*(x, v_k))$$

and

$$f_* : T_x\mathbb{R}^n \to T_y\mathbb{R}^m$$

where $f(x) = y$

Example 136

$$(df_p)^*(w_q) = (f^*w)_p = f^*w(p)((p; v_1), \ldots, (p; v_k)) = w(f(p))(df_p(p; v_1), \ldots, df_p(p, v_k))$$

Example 137
Suppose $q = f(p)$ and $\phi \in \Omega^0(V)$ then

$$f^*\phi(p) = (\phi \circ f)_p = \phi(f(p)) = \phi(q)$$

with reference to 136, this is basically the case where don't have the second bracket containing the pairs in the tangent space.Also

$$(f^*d\phi)_p = (df_p)^*(d\phi)_q = d\phi_p \circ df_p = d(\phi \circ f)_p$$

since

$$df_p : T_q\mathbb{R}^n \to T_p\mathbb{R}^m \quad \text{and} \quad d\phi_p : T_p\mathbb{R}^n \to T_c\mathbb{R} = \mathbb{R} \in \Omega^1(V)$$

This is essentially the case of the second bracket containing the pairs where $c = \phi(q)$ therefore $f^*d\phi = df^*\phi$ and $f^*\phi = \phi \circ f$

Theorem 138
Suppose $\mu \in \Omega^k(V)$ and $\nu \in \Omega^\ell(V)$ then

$$f^*(\mu \wedge \nu) = f^*\mu \wedge f^*\nu$$

Proof. Suppose $q = f(p)$ by definition we have

$$(f^*(\mu \wedge \nu))_p = (df_p)^*(\mu_q \wedge \nu_q) = (df_p)^*\mu_q \wedge (df_p)^*\nu_q = (f^*\mu)_p \wedge (f^*\nu)_p$$

it also helps to visualize the decomposition like so

$$\mu_q \wedge \nu_q \in \Lambda^{k+\ell}(T_q^*\mathbb{R}^n), \quad \mu_q \in \Lambda^k(T_q^*\mathbb{R}^n), \quad \nu_q \in \Lambda^\ell(T_q^*\mathbb{R}^n)$$

$$df_p^*(\mu_q \wedge \nu_q) \in \Lambda^{k+\ell}(T_p^*\mathbb{R}^m), \quad df_p^*(\mu_q) \in \Lambda^k(T_p^*\mathbb{R}^m), \quad df_p^*(\nu_q) \in \Lambda^\ell(T_p^*\mathbb{R}^m)$$

Theorem 139
Let $w \in \Omega^k(V)$ then

$$df^*w = f^*dw$$

Proof. Let $w = \sum a_I dx_I = \sum a_I \wedge dx_I$ then

$$f^*w = \sum f^*a_I \wedge f^*dx_{i_1} \wedge \ldots \wedge f^*dx_{i_k} = \sum f^*a_I e f^*dx_{i_1} \wedge \ldots \wedge f^*dx_{i_k}$$

so since $f^*a_I = a_I \circ f \in \Omega^0(V)$

$$df^*w = \sum df^*a_I \wedge df_{i_1} \wedge \ldots \wedge df_{i_k} + \underbrace{\sum f^*a_I \wedge d(df_{i_1} \wedge \ldots \wedge df_{i_k})}_{=0}$$

where we have used the fact that $f^*dx_i = d(f)^*df(x_i) = d(x_i \circ f) = df_i$. Hence

$$df^*w = \sum_I df^*a_I \wedge f^*(dx_{i_1} \wedge \ldots \wedge dx_{i_k}) = f^*(\sum da_I \wedge dx_I) = f^*dw$$

Theorem 140
For all $w \in \Omega^k(W)$

$$f^*g^*w = (g \circ f)^*w$$

Proof. Let $f(p) = q$ and $g(q) = w$ Then we have...to be continued

4 Integration with differential forms

Definition 141

Let U be an open set in $\mathbb{R}^n$ and let $w \in \Omega^k(U)$ be a differential k-form The **support** of w is

$$\operatorname{supp} w = \overline{\{p \in U : w_p \neq 0\}}$$

The function f is **compactly supported** if $\operatorname{supp} w$ is compact. We define

$$\Omega_c^k(U) = \text{space of all compactly supported k-forms}$$

Given $w \in \Omega_c^n(U)$ we can write

$$w = \phi(x) dx_1 \wedge \ldots \wedge dx_n$$

where $\phi \in \mathcal{C}_0^\infty(U)$

$$\int_U w \equiv \int_U \phi = \int_U \phi(x) dx_1 \ldots dx_n$$

Note that

$$\Omega_c^0(U) = \mathcal{C}_0^\infty(\mathbb{R}^n)$$

where recall the RHS is the set of all compactly supported $\mathcal{C}^\infty$ functions on $\mathbb{R}^n$

Definition 142

Let U, V be open sets in $\mathbb{R}^n$. Let $f : U \to V$ be a $\mathcal{C}^\infty$ diffeomorphism. That is for every $p \in U$, $Df(p) : \mathbb{R}^n \to \mathbb{R}^n$ is bijective. We associate $Df(p)$ with the matrix

$$Df(p) \simeq \left[\frac{\partial f_i}{\partial x_i}(p)\right]$$

The map f is **orientation preserving** if $\det > 0$ everywhere and **orientation reversing** if $\det < 0$ everywhere.

4.1 Sard Theorem

Theorem 143 (Sard theorem)

Let U be open in $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^n$ be a $\mathcal{C}^1(U)$ map. For every $p \in U$ we have the map $Df(p) : \mathbb{R}^n \to \mathbb{R}^n$. We say that p is a **critical point** of f is $Df(p)$ is not *surjective*(in our case not bijective since same dimension mapping).

$$C_f = \text{the set of all critical points of } f$$

Sard theorem then states the image $f(C_f)$ is of measure zero

Proof. Cover U with set of open rectangles $\cup Q_i$ (recall `rudin` that every subset of a metric space has an open cover). Then first consider one such rectangle $Q_i = R$ with width $\frac{\ell}{N} < \delta$ Use continuity to get

$$|Df(x) - Df(y)| \leq \frac{\varepsilon}{n}$$

using $|x - y| < \delta$ and

$$f_i(y) - f_i(x) = Df_i(z)(y - x)$$

(mean value theorem) we get

$$f_i(y) - f_i(x) - Df_i(x)(y-x) = (Df_i(z) - Df_i(x))(x-y)$$

so we have taking sup of both sides

$$(1): \quad |f(y) - f(x) - Df(x)(y-x)| \le n\,|Df(z) - Df(x)|\,|y-x| \le \varepsilon\,|x-y|$$

doing for the mean value theorem too we get

$$(2): \quad |f(y) - f(x)| < c|y-x|$$

where $c = \sup_R Df(x)$ which clearly exists as $Df(R)$ is continuous and R is bounded. From (1) we gave some

$$\left| y - y_0 - A\frac{\ell}{N} \right| \le \varepsilon\frac{\ell}{n}$$

where $A = Df(x_0)$,$y = f(x), y_0 = f(x_0)$ fix a $x_0 \in R \cap C_f$ and consider any $x \in R$ Because A is not bijective it means it must be sent to subspace W with at most $\dim W = n-1$. Because of our result here we know that euclidean distance between $y - y_0 \in V$ and $A\frac{\ell}{N}$ must satisfy our bound of $\varepsilon\frac{\ell}{n}$. That is to say consider unit vector v that is perpendicular to W. Then we must have

$$|a| = |\langle y - y_0, v\rangle| \le \sqrt{n}\frac{\ell}{N}\varepsilon$$

Because unlike other components of $y - y_0$ this is a component which A definitely won't have so it cant just cancel under linear addition of same components. This component of $(y - y_0) \cdot v$ will contribute independently to the lenghth difference. So consider the orthognal decomposition of $y - y_0$

$$y - y_0 = av + w, w \in W$$

We also know from (2) we must have

$$(3): \quad |y - y_0| < c\frac{\ell}{n}$$

Combining everything we see that

$$||w|| \le |a| + ||y - y_0|| \le \sqrt{n}\frac{\ell}{N}(c + \varepsilon) \le 2\sqrt{n}\frac{\ell}{N}c$$

Therefore for any y_0 in W, the ys mapped from the other $x \in R$ traces a rectangle around y_0 in W. We let this be the base of our rectangle traced. The height of our rectangle traced by y around y_0 is clearly given by $|a|$. Hence the bound on total volume of the rectangle traced is

$$(4c)^{n-1}(\sqrt{n}\ell)^n\frac{2\varepsilon}{N^n}$$

because ε can be made arbituarly small, so can this volume as the rest are just constant variables. This is why we cant just take the volume bound to be $(c\frac{\ell}{n})^n$ from (3) because we have to bring in ε to make the volume arbitrarily small. Therefore we introduced it in as a height component in our calculation.

$$f(C_f) = \bigcup f(Q_i \cap C_f)$$

which is a countable union of measure zero sets(as we can make each rectangle Q_i arbituarly small), the result is a measure zero set as desired.

4.2 Poincare lemma

Theorem 144 (Poincare Lemma for open rectangles)
Let U be a connected open subset of $\mathbb{R}^n$ and let $w \in \Omega_c^n(U)$. The following conditions are equivalent

1. $\int_U w = 0$
2. $w = d\mu$ for some $\mu \in \Omega_c^{n-1}(U)$

Let $U = \operatorname{Int} Q$ where $Q = [a_1, b_1] \times \ldots \times [a_n, b_n]$ is a rectangle.

Proof. First we show (2) implies (1) Define

$$\mu = \sum_i f_i dx_1 \wedge \ldots \widehat{dx_i} \ldots \wedge dx_n$$

where $\hat{dx_i} = d_{i-1} \wedge d_{i+1}$ that is every summand above will exlcude dx_i. $\mu \in \Omega_c^{n-1}(U)$. Then

$$w = d\mu = \sum_i \sum_j \frac{\partial f_i}{\partial x_j} dx_j \wedge dx_1 \wedge \ldots \widehat{dx_i} \ldots \wedge dx_n$$

however when $i \neq j$ it means j is equal to one of the elements in $\{i\} = [1, n]/\{j\}$ so we clearly have redundant tensors which we know have a zero tensor product. Hence we can simplify the above to

$$w = d\mu = \sum_i \frac{\partial f_i}{\partial x_i} dx_i \wedge dx_1 \wedge \ldots \widehat{dx_i} \ldots \wedge dx_n$$

which are the cases when $i = j$ so $[1, n] = \{i\} + \{j\}$.

$$w = d\mu = \sum_i (-1)^{i-1} \frac{\partial f_i}{\partial x_i} dx_1 \wedge \ldots \wedge dx_n$$

This follows because to move dx_i from position 1 to its correct position i, it requires $i-1$ swaps. Taking the integral like how we defined in 141 we have

$$\int_U w = \int_U d\mu = \sum (-1)^{i-1} \int_U \frac{\partial f_i}{\partial x_i}$$

where the ability to swap the sum and the integral is due to the fact that reinmann integrals can integrated term by term recall `rudin`. However notice that

$$\int_{a_i}^{b_i} \frac{\partial f_i}{\partial x_i} dx_i = f_i(x)|_{x_i=a_i}^{x_i=b_i} = 0 - 0 = 0$$

because the boundary points of Q are outside of U and that f is compactly supported on U meaning the set of points for which $f \neq 0$ is a closed subset within U, this follows. Then we can also conclude

$$\int_U w = \int_U \frac{\partial f_i}{\partial x_i} = 0$$

by fubini theorem. Just consider

$$\int_{a_1}^{b_1} \ldots \int_{a_i}^{b_i} \ldots \int_{a_n}^{b_n} \frac{\partial f_n}{\partial x_n} dx_n = \int_U \frac{\partial f_i}{\partial x_i}$$

Now we aim the prove the other direction that (1) implies (2). Consider the following lemma.

Lemma 145

Let $w \in \Omega^n(U)$ where $U = Q$ and assume that

$$\int_U w = 0$$

Then for all $1 \leq k \leq n+1$ there exists $\mu \in \Omega_c^{n-1}(U)$ and $f \in \mathcal{C}_0^{\infty}(U)$ such that for a fixed w

$$w = d\mu + f\,dx_1 \wedge \ldots \wedge dx_n$$

and

$$\int f(x_1, \ldots, x_n) \underbrace{dx_k \ldots dx_n}_{dx_i, i \in [k,n]} = 0$$

Proof. Consider the base cases,when $k = 1$

$$\int f(x_1, \ldots, x_n) dx_1 \ldots dx_n = 0$$

Then inferring that $\int (f(x_1, \ldots, x_n))\, dx_1 \ldots dx_n = 0$ we have:

$$0 = \int w = \int d\mu + \int f\,dx_1 \ldots dx_n = \int d\mu + 0$$

Consider the other base case when $k = n+1$

$$\int f(x_1, \ldots, x_n) dx_{n+1} = (f(x_1, \ldots, x_n)) \int dx_{n+1} = x_{n+1}\,(f(x_1, \ldots, x_n)) = 0$$

Then inferring that $(f(x_1, \ldots, x_n)) = 0$ we have:

$$w = d\mu + 0$$

Therefore if we can prove by induction from $k : 1 \to k+1$ then we have $(1) \int w = \int d\mu = 0$ implies $(2) w = d\mu$ as desired.

For the induction step we essentially need to show case k implies $k+1$. That is we want to show there exists $f_{\text{new}} \in \mathcal{C}_0^{\infty}(U)$ and $\mu_{\text{new}} \in \Omega_c^{n-1}(U)$ where

$$w = d\mu_{\text{new}} + f_{\text{new}} dx_1 \wedge \ldots \wedge dx_n$$

and

$$\int f_{\text{new}}(x_1, \ldots, x_n) dx_{k+1} \ldots dx_n = 0$$

using our induction hypothesis. Noticing that

$$g(x_1, \ldots, x_k) = \int f(x_1, \ldots, x_n) d_{k+1} \ldots dx_n$$

since if you integrate a function with respect to $x_{k+1} \to x_n$ that the resultant function will obviously be in the form of g recalling elementary calculus knowledge. Therefore a straightforward candidate for f_{new} is

$$\begin{aligned}
\int f_{\text{new}}(x_1, \ldots, x_n) dx_{k+1} \ldots dx_n &= \int f(x_1, \ldots, x_n) d_{k+1} \ldots dx_n - g(x_1, \ldots, x_k) \int p(x_{k+1}, \ldots, x_n) d_{k+1} \ldots dx_n \\
&= g(x_1, \ldots, x_k) - g(x_1, \ldots, x_k)(1) \\
&= 0
\end{aligned}$$

Where we know there certainly exists $\int p(x_{k+1}, \ldots, x_n) d_{k+1} \ldots dx_n = 1$ due to **partitions of unity** 60. In other words

$$f_{\text{new}} = f - gp$$

Since w is fixed reverse engineering to find $d\mu_{\text{new}}$ we have

$$\begin{aligned} w &= d\mu_{\text{new}} + f_{\text{new}} dx_1 \wedge \ldots \wedge dx_n \\ &= d\mu + (gp + (f - gp)) dx_1 \wedge \ldots \wedge dx_n \\ &= d\mu + f dx_1 \wedge \ldots \wedge dx_n \end{aligned}$$

Hence it is clear we must have

$$d\mu_{\text{new}} = d\mu + (gp) dx_1 \wedge \ldots \wedge dx_n$$

let $-(gp)dx_1 \wedge \ldots \wedge dx_n = d\nu$. Also notice from the by induction hypothesis assumptions we have

$$\int_{a_k}^{b_k} g(x_1, \ldots, x_{k-1}, s) ds = \int f(x_1, \ldots, x_n) dx_k \ldots dx_n = 0$$

To see this just consider fubini theorem that

$$\int_{a_k}^{b_k} g(x_1, \ldots, x_{k-1}, s) ds = \int_{a_k}^{b_k} \left(\int f(x_1, \ldots, x_n) dx_{k+1} \ldots dx_n \right) ds$$

where $ds = dx_k$ clearly. Define

$$h(x_1, \ldots, x_k) = \int_{a_k}^{x_k} g(x_1, \ldots, x_{k-1}, s) ds$$

Following the intuitition behind to proof of (2) implying (1) we similary define

$$\nu = (-1)^k h(x_1, \ldots, x_k) p(x_{k+1}, \ldots x_n) dx_1 \wedge \ldots \widehat{dx_k} \ldots dx_n$$

so we have

$$\begin{aligned} d\nu &= (-1)^k \sum_j \frac{\partial}{\partial x_j}(hp) dx_j \wedge dx_1 \wedge \ldots \widehat{dx_k} \ldots dx_n && (1) \\ &= (-1)^k \frac{\partial}{\partial x_k}(h) p dx_k \wedge dx_1 \wedge \ldots \widehat{dx_k} \ldots dx_n && (2) \\ &= (-1) \frac{\partial}{\partial x_k}(h) p dx_k \wedge dx_1 \wedge \ldots \widehat{dx_k} \ldots dx_n && (3) \end{aligned}$$

where recall if $j \neq k$ the summand is zero and that we require $k-1$ swaps to bring dx_k to the right position. Now let us prove that everything is still compactly supported as required. Suppose μ, f supported on $[c_1, d_1] \times \ldots [c_k, d_k]$ Consider

$$h(x) = \int_{a_k}^{b_k} g(s) ds - \int_{x_k}^{b_k} g(s) ds$$

where $a_k < c_k < d_k < b_k, x_k > d_k$ and $\operatorname{supp} g \subseteq [c_k, d_k]$ since the integrand on the RHS in the definition of g disappears outside this partition. Therefore the first integral on the RHS of $h(x)$ disappears while $g(s) = 0$ outside of its support so the 2nd RHS integral also zero. So outside of $[c, d]$ we have shown $h = 0$ so it is also supported on $[c_1, d_1] \times \ldots [c_k, d_k]$. We don't have to check for p has we know partitions of unity is compactly supported. Now we need check that our functions are C^∞. Again by partition of unity definitions we don't need to check for p. For g we know that is an integral of f so it is simply a lower order differential of f.Because $f \in C^\infty$ then clearly so is g. It is also clear from our formulations that $\nu \in \Omega_c^{n-1}(U)$ so that $\mu_{\text{new}} = \mu - \nu \in \Omega_c^{n-1}(U)$ as desired. $\square$

Remark 146. *Note that the n-form $w \in \Omega_c^n(U)$ which is the space of all n forms of class C^∞ compactly supported on U. We can actually relax the condition a bit by requiring w to be of class C^r where $r \geq 1$(is still compactly supported on U though). Notice in our derivation we had to differentiate the scalar field in our form f once(look at the first two lines in the proof).*

We now aim to prove the more general case by reducing it to the rectangle case. Again we really only require $w \in C^r$ where $r \geq 1$ but must still be compactly supported.

Corollary 147 (General Poincare Lemma)

Let U be a *connected* open subset of $\mathbb{R}^n$ and let $w \in \Omega_c^n(U)$. The following conditions are equivalent

1. $\int_U w = 0$
2. $w = d\mu$ for some $\mu \in \Omega_c^{n-1}(U)$

Proof. To prove (2) immplies (1) define a family of recantagles $\{R_i, i \in \mathbb{N}\}$ where

$$U = \bigcup \operatorname{Int} R_i$$

Since the support of μ is cmpact, it is covered by finitely many of such rectangles. So take a partition of unity $\{\phi_i\}$ subordinate to $\{R_i\}$. Then there exists a real number N where

$$\underbrace{\mu}_{\text{supported on U}} = \sum_{i=1}^{N} \underbrace{\phi_i \mu}_{\text{supported on } \operatorname{Int} R_i}$$

knowing that $\phi \in \mathcal{C}^\infty$ and that differential operators are linear on zero forms and the fact that riemann integrals can be done term by term we have

$$\int d\mu = \sum_i \int d(\phi_i \mu)$$

where we have brought the sum both out of the differential and the integral. From poincare lemma proof for rectangles since $\phi_i \mu$ which is compact supported on a rectangle we have every $\int d(\phi_i \mu) = 0$ so we have $\int d\mu = \int d(\sum \phi_i \mu) = 0$ as desired. To prove that (2) implies (1) consider that this is equivalent to showing if

$$\int w_1 = \int w_2$$

then $\int w_1 - w_2 = 0 \Rightarrow d\mu = w_1 - w_2$. We denote this relationship by the equivalence relation $w_1 \sim w_2$. We first prove the lemma below before we proceed

Lemma 148

Consider sequence of rectangles R_i such that $\operatorname{supp} \phi_i w \subset \operatorname{Int} R_i$. Due to connectedness, for any fixed R_0 and $x \in R_n$ there exists a finite chain of rectangles $R_0, \ldots, R_n$ that connect R_0 and R_N has no "breakages". That is $(\operatorname{Int} R_i) \cap (\operatorname{Int} R_{i+1})$ is non empty

Proof. Just refer to 227 □

Now back to our general poincare lemma proof, Construct forms ν_i such that $\operatorname{supp} \nu_i \subseteq \operatorname{Int} R_i \cap \operatorname{Int} R_{i+1}$ and that $\int \nu_i = 1$, again made possible by partitions of unity. Therefore clearly

$$\operatorname{supp}(\nu_i - \nu_{i+1}) \subseteq \operatorname{Int} R_{i+1}$$

so on this rectangle Int R_{i+1} by our previous poincare lemma,

$$(\int \nu_i - \nu_{i+1}) = 0 \Rightarrow \nu_i \sim \nu_{i+1}$$

Now notice that

$$c_i = \int \underbrace{\phi_i w}_{\text{supported in some } R_i} = \int \underbrace{c_i w_0}_{\text{supported in some } R_0} = c_i$$

where we define $\int w_0 = 1$ and $\int \phi_i w = c_i$. We can fix $c_i w_0$ to be supported by some R_0 because c_i is just some constant, so the support is completely dependent on w_0 which can be made by partitions of unity where we decide which rectangle it would be supported in. Therefore knowing that R_i and R_0 can be connected by a chain of rectangles with no breakages by construction of ν_i we have a chain of equivalence relations

$$c_i w_0 \sim c_i \nu_0 \sim c_i \nu_1 \ldots \sim c_i \nu_N \sim \phi_i w$$

Since we have proven $\int \phi_i w = \int c_i w_0 \Rightarrow c_i w_0 \sim \phi_i w$ we have that (1) implies (2) as desired. □

4.3 Proper Maps and Degree

We now introduce a class of functions that remain compactly supported under the pullback operation

Definition 149

Let $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^k$ and let $f : U \to V$ be a continuous map. The map f is **proper** if for all compact subsets $K \subseteq V$ the set $f^{-1}(K)$ is compact

Let $w \in \Omega^k(V)$. Define the map

$$f^* : \Omega^k(V) \to \Omega^k(U)$$

meaning that

$$w = g(y_1, \ldots, y_n) dy_{i_1} \wedge \ldots \wedge dy_{i_k}$$

and

$$f^* w = g(f(x)) df_{i_1} \wedge \ldots \wedge df_{i_k}$$

So if f is proper and $w \in \Omega^n_c(V)$ then $\text{supp}(f^* w)$ is compact

$$f^{-1}(\text{supp } w) \supseteq \text{supp}(f^* w)$$

$w \in \Omega^n_c(V)$ which implies $f^* w \in \Omega^n_c(U)$. Therefore the integrals defined by

$$\int_U f^* w = (\deg f) \int_V w$$

here makes sense. However this is not the only condition where $\int f$ makes sense. As we will show below, we can use **Poincare lemma** to define such integrals too.

Definition 150

Let U, V to each be a connected open subsets of $\mathbb{R}^n$ and let $f : U \to V$ be a $\mathcal{C}^\infty$ map then for **all** $w \in \Omega_c^n(V)$ we have

$$\int_U f^* w = (\deg f) \int_V w$$

We say all because our definition does not depend on choice of w since we defined

$$\deg f \equiv \int f^* w_0$$

where $w_0 \in \Omega_c^n(V)$ and

$$\int w_0 = 1$$

which is possible by partitions of unity. Show that this is well defined by Poincare lemma

Proof. and let

$$\int w = c$$

so like how we did before we do

$$\int w = \int c w_0$$

since V is open connected by the generalized poincare lemma we have $w \sim c w_0$ where $w = c w_0 + d\mu$ Then

$$f^* w = f^*(c w_0) + f^*(d\mu) = f^*(c w_0) + d(f^* \mu)$$

Again by poincare lemma this shows $f^* \sim f^*(c w_0)$ Now taking integrals on each side by linearity we have

$$\begin{aligned} \int f^* w &= c \int f^*(w_0) \\ &= c \deg f \\ &= \left(\int w \right) \deg f \end{aligned}$$

5 general degree counting and applications

Theorem 151

Let U, V be connected open sets in $\mathbb{R}^n$ and $f : U \to V$ be a proper C^∞ map. The degree of f is

$$\deg(f) = \sum_{i=1}^{N} \sigma_{p_i}$$

where

$$\sigma_{p_i} = \begin{cases} +1 & Df(p_i) \text{ is orientation preserving} \\ -1 & Df(p_i) \text{ is orientation reserving} \end{cases}$$

Proof. Select a regular value $q \in V - f(C_f)$ meaning for any $p \in f^{-1}(q) \Rightarrow p \notin C_f$. By the **inverse function theorem** the map f is a diffeomorphism of a neighbourhood of U_p of $p \in U$ onto a neighbourhood of q. In particular since f is bijective for every p,

$$U_p \cap f^{-1}(q) = \{p\}$$

Therefore we know each U_p covers its own p only and it forms an open cover of $f^{-1}(q)$. Because f is a proper map and q is just one regular value, so it is clearly a compact set that means $f^{-1}(q)$ is compact as well. Then we know there exists a finite subcover of $\{U_{p_i} : i = 1, \ldots, N\}$ of $f^{-1}(q)$. Now first choose some neighbourhood V_0 of q such that $f^{-1}(V_0) \subseteq \bigcup U_{p_i}$. Then replace each U_{p_i} with $U_{p_i} \cap f^{-1}(V_0)$. So we have

1. f is a diffeomorphism of U_{p_i} onto V_0
2. ${V_0}^{-1} = \bigcup U_{p_i}$
3. U_{p_i} are disjoint from each other

Now choose $w \in \Omega_c^n(V_0)$ where $\int_V w = 1$(recall possible by partitions of unity). Then we we have

$$\begin{aligned} \int_U f^* w &= \sum_i \int_{U_{p_i}} f^* w && (1) \\ &= \sum_i \sigma_{p_i} \int_{V_0} w && (2) \\ &= \sum_i \sigma_{p_i} && (3) \end{aligned}$$

where (1) follows because out side V_0 the integral must vanish as the form w compactly supported on V_0 and $\sigma_{p_i} = \pm 1$ (see below on change of variables for more). But we know

$$\int_U f^* w = \deg f \int_U w = \deg f$$

this makes sense by poincare lemma as U, V connected. So therefore we have

$$\sum \sigma_{p_i} = \deg f$$

as desired. Moreover the choice of regular value should not matter because for the same $f : U \to V$, $\deg f$ will always be the same as shown in its defintion earlier.

Corollary 152

if $f : U \to V$ isn't onto then $\deg f = 0$

Proof. If not onto there exists $q \in V$ but not in the image of f. This means that q is a regular value as it certainly belongs to $V - f(C_f)$. Then using the method of calculations we had above, knowing that $f^{-1}(q)$ is empty(not in U for sure) we know that $\deg f = 0$

Corollary 153

If $\deg f \neq 0$ then $f : U \to V$ is onto.

Proof. it is the contrapositive of the above.

Theorem 154 (Brouswer fixed point theorem)
Let B^n be the closed unit ball in $\mathbb{R}^n$

$$\{x \in \mathbb{R}^n, ||x|| \leq 1\}$$

If $f : B^n \to B^n$ is a C^2 mapping then f has a fixed point. That is f maps some point $x_0 \in B^n$ onto itself. That is

$$f(x_0) = x_9$$

Proof. Is another long topological proof...

6 change of variables formula

6.1 preliminaries

Fact 155
Relating to change of coordinates(see change of variables formula for more)

$$\int g(f(x)) \det \left[Df \right] dx_1 \wedge \ldots \wedge dx_n = (\deg f) \int g(y_1, \ldots, y_n) dy_1 \wedge \ldots \wedge dy_n$$

Theorem 156

$$\deg(f \circ g) = (\deg g)(\deg f)$$

hint: recall $(f \circ g)^* = g^* \circ f^*$

Proposition 157
Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a translation $f(x) = x + a$. Show that $\deg(f) = 1$

Proof. Consider the effect of f on some form $w = \phi(x) dx_1 \wedge \ldots \wedge dx_n$
Consider the case when $n = 1$

$$\int w(x)dx = \int w(x - a)dx = \int f^* w(x)dx$$

which can easily be seen by change of variables $t = t - a$ We can clearly extend this to multiple variable integration too by fubini. That is

$$\int w(x_1, \ldots, x_n) dx_1, \ldots, dx_n = \int w(x_1 - a_1, \ldots, x_n - a_n) dx_1, \ldots, dx_n$$

Proposition 158
Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be the mapping

$$f(x_1, \ldots, x_n) = (x_1 + \lambda x_2, x_2, \ldots, x_n)$$

Proof. This is similar to the above as in the x_1 slot x_2 is simply a constant when integrating with respect to x_1. Now we extend to multivariables via fubini once more.

Proposition 159

Let σ be a permutation of the numbers $1, \ldots, n$ and let $f_\sigma : \mathbb{R}^n \to \mathbb{R}^n$ be the diffeomorphism

$$f_\sigma(x_1, \ldots, x_n) = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

Prove that $\deg f_\sigma = \operatorname{sgn}(\sigma)$

Proof. Looking at the form $w = \phi(x) dx_1 \wedge \ldots \wedge dx_n$ it is easy to see $f^* w = \operatorname{sgn}(\sigma) w$

Proposition 160

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be the mapping

$$f(x_1, \ldots, x_n) = (\lambda x_1, x_2, \ldots, x_n)$$

with $\lambda \neq 0$. Show that $\deg f = +1$ if λ is positive and -1 if λ is negative

Proof. Let

$$\omega = g(x_1, \ldots, x_n)\, dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n,$$

then the pullback:

$$f^*\omega = g(\lambda x_1, x_2, \ldots, x_n)\, d(\lambda x_1) \wedge dx_2 \wedge \cdots \wedge dx_n.$$

We want to compute the integral:

$$\int_U f^*\omega = \int_U \lambda\, g(\lambda x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \cdots dx_n.$$

To simplify the integral, perform a change of variables:

$$u_1 = \lambda x_1, \quad u_i = x_i \text{ for } i = 2, 3, \ldots, n.$$

Then, $du_1 = \lambda\, dx_1$, and $du_i = dx_i$ for $i = 2, 3, \ldots, n$. Substituting into the integral:

$$\int_U f^*\omega = \lambda \int_{f(U)} g(u_1, u_2, \ldots, u_n) \cdot \frac{1}{\lambda}\, du_1 du_2 \cdots du_n.$$

The factors of λ cancel out so we obtain

$$\int_U f^*\omega = \int_{f(U)} \omega.$$

Corollary 161

Let A be a non-singular $n \times n$ matrix and $f_A : \mathbb{R}^n \to \mathbb{R}^n$ the linear mapping associated with A. Then $\deg(f_a) \pm 1$ corresponding to the polarity of $\det(A)$

Consider that f_A can be writtten as

$$f_A = f_{E_1} \circ \ldots \circ f_{E_k}$$

where E_i are elementary matrices. Clearly all the propositions above describe the actions of elementary matrix operations that is if you recall

1. (row swap) $\det A' = -\det A$
2. (scalar multiplication row j) $\det A' = c \det A$
3. (add multiple row i to row j) $\det A' = \det A$

and

$$\deg(f_A) = \deg(f_{E_1}) \dots \deg(f_{E_k}) \quad \text{and} \quad \det(A) = \det(E_1) \dots \det(E_k)$$

Clearly the polarity corresponds, but not the magnitude due to (2).

6.2 change of coor and variables formula

Theorem 162 (Change of coordinates formula)
Given $w = \phi(x) dx_1 \wedge \dots \wedge x_n$ where $\phi \in \mathcal{C}^\infty$

$$f^* w = \phi(f(x)) \det \left[\frac{\partial f_i}{\partial x_j} \right] dx_1 \wedge \dots \wedge dx_n$$

Proof. Suppose that

$$f : U \to V$$

be a C^∞ map. So we have

$$(df)_p : T_p U \to T_q V$$

$$(df^*)_p : \Lambda^n(T_q^* V) \to \Lambda^n(T_p^* U)$$

Consider $(dx_1)_p \wedge \dots (dx_n)_p \in \Lambda^n(T_p^* U)$ where $((dx_1)_p, \dots, (dx_n)_p)$ is basis of $T_p^* U$
Consider $(dy_1)_q \wedge \dots (dy_n)_q \in \Lambda^n(T_p^* U)$ where $((dy_1)_q, \dots, (dy_n)_q)$ is basis of $T_q^* V$ so expressing change of variables using the linear map f we have

$$f(x_i) = \sum_j a_{i,j} y_j$$

and

$$df_p^* (dy_j)_q = \sum_k a_{jk} (dx_k)_p$$

Then

$$\begin{aligned}
df_p^*(\phi(p)(dy_1)_q \wedge \dots \wedge (dy_n)_q) &= \phi(f(p)) df_p^*(dy_1)_q \wedge \dots \wedge df_p^*(dy_n)_q && (1)\\
&= \phi(f(p)) \left(\sum_{k_1}^n a_{1,k_1} (dx_{k_1})_p \right) \wedge \dots \wedge \left(\sum_{k_n}^n a_{n,k_n} (dx_{k_n})_p \right) && (2)\\
&= \phi(f(p)) \sum_{k_1,\dots,k_n} (a_{1,k_1} \dots a_{n,k_n})((dx_{k_1})_p \wedge \dots \wedge (dx_{k_n})_p) && (3)\\
&= \phi(f(p)) \sum_\sigma (a_{1,\sigma(1)} \dots a_{n,\sigma(n)})((dx_{\sigma(1)})_p \wedge \dots \wedge (dx_{\sigma(n)})_p) && (4)\\
&= \phi(f(p)) \sum_\sigma (-1)^\sigma (a_{1,\sigma(1)} \dots a_{n,\sigma(n)})((dx_1)_p \wedge \dots \wedge (dx_n)_p) && (5)
\end{aligned}$$

recall (4) follows because the only non-zero terms non-repeating wedge products so we tae the permutations only and then $(-1)^\sigma$ is the number of swaps to order a permutation to the correct ascending order like so. However we know that letting $q = f(p)$ we have

$$\begin{aligned}
(df)^*_p(dy_i)_q = (f^* dy_i)_p = (dy_i \circ f)_p = (df_i)_p = (f^*(dy_i))(p)(p;v) &= dy_i(q)(f_*(p;v)) && (6)\\
&= dy_i(q)(q; Df(p)\cdot v) && (7)\\
&= \text{row ith component of } Df(x)\cdot v && (8)\\
&= Df_i(p)\cdot v && (9)\\
&= \sum_j (D_j f_i(p))\cdot(v_j) && (10)\\
&= \sum_j \left(\frac{\partial f_i}{\partial x_j}\right)\cdot(dx_j(p)(p;v)) && (11)\\
&= \sum_j \frac{\partial f_i}{\partial x_j}(p)(dx_j)_p && (12)
\end{aligned}$$

(8) follows from 122. So it follows that

$$A = [a_{i,j}] = \left[\frac{\partial f_i}{\partial x_j}(p)\right]$$

Also we have seen this result before from 8, so we have just proven it again. Finally notice that

$$\sum_\sigma (-1)^\sigma (a_{1,\sigma(1)} \dots a_{n,\sigma(n)}) = \det A$$

by definition 117 □

Lets see how this makes sense on maps from $\mathbb{R}^k \to \mathbb{R}^n$

Corollary 163

Let A be open in $\mathbb{R}^k$ and let $a : A \to \mathbb{R}^n$ be a C^∞ map. Let x the general point of $\mathbb{R}^k$ and let y denote the general point of $\mathbb{R}^n$. THen dx_i and dy_i denote the elementary 1-forms of $\mathbb{R}^k$ and $\mathbb{R}^n$ respectively.

1. $a^*(dy_i) = da_i$
2. If $I = (i_1, \dots, i_k)$ is an ascending k-tuple from the set $\{1, \dots, n\}$ then

$$a^*(dy_I) = \det \frac{\partial a_I}{\partial x} dx_1 \wedge \dots \wedge dx_k$$

where

$$\frac{\partial a_I}{\partial x} = \frac{\partial(a_{i_1}, \dots, a_{i_k})}{\partial(x_1, \dots, x_k)} = \left[\frac{\partial a_{i_p}}{\partial x_q}\right]$$

Proof. Simply consider from the above

$$\begin{aligned}
df_p^*(\phi(p)(dy_{\mathbf{1}})_q \wedge \ldots \wedge (dy_{\mathbf{k}})_q) &= \phi(f(p))df_p^*(dy_{\mathbf{1}})_q \wedge \ldots \wedge df_p^*(dy_{\mathbf{k}})_q && (13)\\
&= \phi(f(p))\left(\sum_{k_1}^{k} a_{\mathbf{1},k_1}(dx_{k_1})_p\right) \wedge \ldots \wedge \left(\sum_{k_{\mathbf{k}}}^{k} a_{\mathbf{k},k_{\mathbf{k}}}(dx_{k_{\mathbf{k}}})_p\right) && (14)\\
&= \phi(f(p)) \sum_{k_1,\ldots,k_{\mathbf{k}}} (a_{\mathbf{1},k_1}\ldots a_{\mathbf{k},k_{\mathbf{k}}})((dx_{k_1})_p \wedge \ldots \wedge (dx_{k_{\mathbf{k}}})_p) && (15)\\
&= \phi(f(p))\sum_{\sigma}(a_{\mathbf{1},\sigma(\mathbf{1})}\ldots a_{\mathbf{k},\sigma(\mathbf{k})})((dx_{\sigma(\mathbf{1})})_p \wedge \ldots \wedge (dx_{\sigma(\mathbf{k})})_p) && (16)\\
&= \phi(f(p))\sum_{\sigma}(-1)^{\sigma}(a_{\mathbf{1},\sigma(\mathbf{1})}\ldots a_{\mathbf{k},\sigma(\mathbf{k})})((dx_{\mathbf{1}})_p \wedge \ldots \wedge (dx_{\mathbf{k}})_p) && (17)
\end{aligned}$$

We let $[\mathbf{1},\ldots,\mathbf{k}] = [i_1,\ldots,i_k]$ be the ascending tuple from the set $\{1,\ldots,n\}$

Definition 164

Let U and V be connected open subsets of $\mathbb{R}^n$. If $f : U \to V$ is a $\mathcal{C}^2$ diffeomorphism, recall then the determinant $Df(x)$ for all $x \in U$ is hence non-zero. We will say f is orientation preserving if the sign of the determinant is positive and orientation reversing if it is negative.

The reason has got to do with

$$(df)^* : \Lambda^n(U)_+ \to \Lambda^n(V)_+$$

Where this is a map between 1-dimensional vector dual spaces so

$$f^*w = \lambda_A = w$$

where $\lambda_A = \det Df(x)$ if you recall. Now we want to prove that

Theorem 165

Let U and V be connectd open subsets of $\mathbb{R}^n$. If $f : U \to V$ is a $\mathcal{C}^2$ diffeomorphism.
The *degree* of f is $+1$ is f is orientation preserving and -1 if f is orientation reversing. Recall our definition of degree here makes sense by 150

Proof. Given a point $a_1 \in U$, let $a_2 = -f(a_1)$ and for $i = 1, 2$ let $g_i : \mathbb{R}^n \to \mathbb{R}^n$ be the translation $g_i(x) = x + a_i$ then recall from previously

$$g_2 \circ f \circ g_1$$

has the same degree as f. Hence redefine f using such a mapping where 0 in the domain of f, $f(0) = 0$ and has a degree equal to the original. Since $f(0) = 0$ and f is a diffeomorphism, we can certainly can associate a bijective linear mapping $A : \mathbb{R}^n \to \mathbb{R}^n$ with it that is

$$A = Df(0)$$

so we have

$$D(f^{-1} \circ f) = Df^{-1} \circ f = A^{-1} \circ f = I$$

Let $f_{\text{new}} = f^{-1} \circ f$ so

$$\deg(f_{\text{new}}) = \operatorname{sgn}(\det A) \deg f = \operatorname{sgn}(\det A^{-1}) \deg f = \deg A \deg(f)$$

since recall `artin algebra 1` $\det A = \frac{1}{\det A^{-1}}$ and $\deg A = \pm 1$ and $\deg f = \pm 1$ we can prove the theorem if we can prove that $\deg f_{\text{new}} = 1$ since that will imply

$$\deg f = \begin{cases} 1 & \deg A = 1 \\ -1 & \deg A = -1 \end{cases}$$

Where we know from previously that $\deg A = \pm 1$ follows $\operatorname{sgn}\det A$ determines orientation preserving and reversing by definition.

Now again replacing f with f_{new} we now consider $g(x) = x - f(x)$ so $Dg(0) = 0$. Notice that all our simplications so far are exactly the same as how we did for *inverse function theorem*. The main reason is to be able to make use of the continuity of $\frac{\partial g_i}{\partial x_j}(0)$ so it is easier for us to set up our topological proof as follows. Roughly, the big idea after this is that after a large chunk of topological proof using partitions of unity to define to some $\tilde{f}$, balls and compact supports etc we can show that there exists

$$\deg f = \int f^* w = \int \tilde{f}^* w = \deg \tilde{f}$$

on a set where both $f, \tilde{f}$ are supported on and

$$\deg \tilde{f} = \int \tilde{f}^* w = \int w = 1$$

on another defined set where $\tilde{f}$ is supported on. And because deg is some integer independent of whatever compact supported form used(by definition) hence we have $\deg f = 1$

Theorem 166 (Change of variables formula)
Let U and V be connectd open subsets of $\mathbb{R}^n$. If $f : U \to V$ is a $\mathcal{C}^2$ diffeomorphism.
Let $\phi : V \to \mathbb{R}$ be a compactly supported continuous function. Then

$$\int_U \phi \circ f(x) \left|\det(Df)(x)\right| dx = \int_V \phi(y) dy$$

Proof. Note that the proof above already satisfies the case for when ϕ is $\mathcal{C}^1$. Recall 144 we only required the form to be at least C^r where $r \geq 1$ in our derivation of poincare lemma which allowed our definition of degree here to make sense.

$$\int_U \phi \circ f(x) \det(Df)(x) dx = \deg f \int_U f^* \phi = \int_V \phi(y) dy$$

for 2nd equality we have proven earlier $\deg f = 1$ and that this integral is well defined by pointcare lemma. The first equality is because of **change of coordinates theorem**. And we don't have to take absolute value because we know $\deg f = 1$ meaning $\det Df(x) > 0$(orientation preserving) So now it remains to prove the case for $\mathcal{C}^0$ (just continuous,not continuously differentiable). Turns out there is still a way to sure the compact support is well defined...but its very topological and I'm short of time :(

Theorem 167
Let U, V be connected open sets in $\mathbb{R}^n$ and let $f : U \to V$ be a proper C^∞ map. Let B be a compact subset of V and let $A = f^{-1}(B)$. If U_0 is an open subset of U with $A \subseteq U_0$ then there exists an open subset V_0 of V with $B \supseteq V_0$ such that $f^{-1}(V_0) \subseteq U_0$

Definition 168
A point $q \in V$ is regular value of f is $q \in V - f(C_f)$

7 other applications

7.1 diffeomorphisms in $\mathbb{R}^n$

Proposition 169
A linear map $h : \mathbb{R}^n \to \mathbb{R}^n$ is a C^∞ diffeomorphism if it is bijective.

Proof. If bijective then the linear map h^{-1} also exists. The proposition follows from the fact that linear maps $\mathbb{R}^n \to \mathbb{R}^n$ are C^∞. Consider we can write

$$h_i(x_1, \ldots, x_n) = a_1 x_1 + \ldots a_n x_1$$

From here we already can see that h_i is continuous. To see this recall from `functional analysis` that a linear operator is continuous if and only there exists $C > 0$ such that $\forall v \in V, ||Tv||_w \leq C\,||v||_V$. Using the standard metric which is just the **euclidean norm** this is evident from our equationn here that bounded pre-images in our domain are sent to bounded images in our codomain. We can also see that $\frac{\partial h_i}{\partial x_j} = a_j$. For higher order derivatives then it will become the zero function which is infinitely differentiable(which implies continuous since derivative exists).This is because although zero, the limit still exists

$$\lim_{h \to 0} \frac{0(x+h) - 0(x+)}{h} = \frac{0}{0} = 0$$

Lemma 170
Let A be open in $\mathbb{R}^n$ let $g : A \to \mathbb{R}^n$ be a function of class $\mathcal{C}^1$. If the subset E of A has measure zero in $\mathbb{R}^n$ then the set $g(E)$ also has measure zero in $\mathbb{R}^n$.

Proof. Study topology first before returning!

Remark 171. *Note that differentiability is needed for this lemma if g is merely continuous then the image of a set of measure zero need not have measure zero. This fact follows from the existence of a continuous map* $f : [0,1] \to [0,1]^2$ *known as the* **Pleano space-filling curve** *which is studied in* `topology`

Theorem 172
Let $g : A \to B$ be a *diffeomorphism* of class $\mathcal{C}^r$ where A and B are open sets in $\mathbb{R}^n$. Let D be a compact subset of A and let $E = g(D)$

(a)
$$g(\operatorname{Int} D) = \operatorname{Int} E \quad \text{and} \quad g(\operatorname{Bd} D) = \operatorname{Bd} E$$

(b) If D is rectifiable so is E

Proof. Since g^{-1} is continuous, recall `rudin` this implies $g(U)$ is open in B for every open set U in A. Thus

$$g(\operatorname{Int} D) \subset g(D) \quad \rightarrow \quad g(\operatorname{Int} D) \subset \operatorname{Int} g(D)$$

since $g(\operatorname{Int} D)$ and $\operatorname{Int} D$ are open and the largest possible open set of $g(D)$ is $\operatorname{Int} g(D)$.Similarly

$$g(\operatorname{Ext} D) \subset (g(D))^c \quad \rightarrow \quad g(\operatorname{Ext} D) \subset \operatorname{Ext} g(D)$$

since $g(\operatorname{Ext} D)$ and $\operatorname{Ext} D$ are open and largest possible open set of $(g(D))^c$ is $\operatorname{Ext} g(D)$. Therefore putting the above 2 relations together we have

$$g(\operatorname{Ext} D \cup \operatorname{Int} D) \subset \operatorname{Ext} g(D) \cup \operatorname{Int} g(D) \quad \rightarrow \quad g(\operatorname{Bd} D) \supset \operatorname{Bd} g(D)$$

which is obtained by taking complements on both sides again recall this relation in `rudin`. Then (a) by substituting $g(D) = E$ For (b), if D is rectifiable than $\operatorname{Bd} D$ has measure zero. Then by the preceding lemma $g(\operatorname{Bd} D)$ has measure zero too. However we also see that by the same reasoning, this applies to the diffeomorphism $g^{-1} : B \to A$ as well in particular

1. $g^{-1}(\operatorname{Int} E) \subset \operatorname{Int} D \quad \rightarrow \quad \operatorname{Int} E \subset g(\operatorname{Int} D)$
2. $g^{-1}(\operatorname{Bd} E) \supset \operatorname{Bd} D \quad \rightarrow \quad \operatorname{Bd} E \supset g(\operatorname{Bd} D)$

TWhere the RHS is obtained by taking g on both sides of (1) and (2) respectively. Therefore the only solution that satisfies both cases is equality.

Definition 173
Let $h : A \to B$ be a diffeomorphism of open sets in $\mathbb{R}^n$ where $n \geq n$,is given by the equation

$$h(x) = (h_1(x), \ldots, h_n(x))$$

Given i we say that h preserves the ith coordinate if $h_i(x) = x_i$ for all $x \in A$. If h preserves the ith coordinate for some i then h is called a **primitive diffeomorphism**

7.2 meaning of determinant

Theorem 174
Let A be an $n \times n$ matrix. Let $h : \mathbb{R}^n \to \mathbb{R}^n$ be the linear transofmration $h(x) = Ax$. Let S be a rectifiable set in $\mathbb{R}^n$ and let $T = h(S)$. Then

$$v(T) = |\det A|\, v(S)$$

Proof. Recall 169 which explains that linear maps between finite dimensional vector spaces are continuous & bounded(that is bounded set is sent to bounded set) and that bijective linear maps are smooth.Hence we know T is bounded as well given that S is bounded.Since smooth in the case when A is non-singular , h is a C^∞ diffeomorphism. Therefore by the previous theorems we know that $h(\operatorname{Int} S) = \operatorname{Int} T$ and that T is rectifiable since S is rectiable. Note that S being compact is not necessary for a well defined integral here recall 50. Therefore

$$v(T) = v(\operatorname{Int} T) = \int_{\operatorname{Int} T} 1 = \int_{\operatorname{Int} S} |\det Dh|$$

this integral is well defined where the last equlaity follows by change of base theorem so $\det A = 0$. We need to show that this implies $v(T) = 0$. Consider the case when A is singular. Then h maps $\mathbb{R}^n$ to some $\mathbb{R}^p$ where $p < n$. However we know that the set $\mathbb{R}_p$ has measure zero($m(v) = 0$) in $\mathbb{R}^n$ since it can be covered by rectangles of this form

$$(x_1, \ldots, x_p, \pm\delta_{p+1}, \ldots, \pm\delta_n)$$

where we can let δ_i be arbitrarily small. Then relating this to *volume* we have $v(T) = \int_T 1 = 0$ since we know the integral vanishes outside the set of measure zero.

7.3 Volume

Theorem 175

There is a unique function V defined by

$$V(x_1, \ldots, x_k) = V(X) = [\det(X^t X)]^{\frac{1}{2}}$$

that assigns to each k-tuple $(x_1, \ldots, x_k)$ of elements in $\mathbb{R}^n$, which we can write as an $n \times k$ matrix X, a non-negative number such that

1. If $h : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal transformation then $V(h(x_1), \ldots, h(x_k)) = V(x_1, \ldots, x_k)$
2. If $y_1, \ldots, y_k$ belong to the subspace $\mathbb{R}^k \times 0$ of $\mathbb{R}^n$ so that

$$y_i = \begin{bmatrix} z_i \\ 0 \end{bmatrix}$$

   for $z_i \in \mathbb{R}^k$ then

$$V(y_1, \ldots, y_k) = |\det[z_1, \ldots, z_k]|$$

Proof. Let

$$V(X) = (F(X))^{\frac{1}{2}}$$

For (1) Let the orthogonal transformation $h : \mathbb{R}^n \to \mathbb{R}^n$ be written as $h(x) = Ax$ where A is an orthogonal $n \times n$ matrix. Then

$$F(AX) = \det((AX)^t(AX)) = \det(X^t X) = \det^2 X = F(X)$$

For (2) consider Let

$$Y = \begin{bmatrix} Z \\ 0 \end{bmatrix}$$

where Z is a $k \times k$ matrix while Y is an $n \times n$ matrix then

$$F(Y) = \det(\begin{bmatrix} Z^t & 0 \end{bmatrix} \begin{bmatrix} Z \\ 0 \end{bmatrix}) = \det(Z^t Z) = \det^2(Z)$$

Uniqueness follows because if Z or X is dependent its determinant vanishes so $V(X), V(Z) = 0$.

Definition 176

Let $k \leq n$. Let A be open in $\mathbb{R}^k$ and let $\alpha : A \to \mathbb{R}^n$ be a map of $\mathcal{C}^r(r \geq 1)$. The set $Y = \alpha(A)$ together with the map α constitute what is called a **parameterized manifold** of dimension k. We denote it by Y_α and define its k-dimensional volume by

$$v(Y_\alpha) = \int_A V(Da)$$

We define the **integral of f over Y_α with respect to volume** by the equation

$$\int_{Y_a} dV = \int_A V(Da) \quad \text{and} \quad \int_{Y_\alpha} f\,dV = \int_A (f \circ \alpha)V(D\alpha)$$

provided this integral exists

Remark 177. *We denote that dV to be the integral with respect to volume is distinct from differential operators dw.*

Example 178

The special case where A is open in $\mathbb{R}^1$ and $a : A \to \mathbb{R}^n$ be the parameterization of our 1-manifold. In that case we denote the **displacement** ds to be the 1 dimensional version of dV since $(Da)^t Da$ is a $k \times k$ matrix and now $k = 1$. that is

$$v(Y_a) = \int_A V(Da)$$

and

$$\int_{Y_a} ds = \int_A V(Da) \quad \text{and} \quad \int_{Y_a} f\,ds = \int_A (f \circ a)V(Da)$$

We now show that our definition is indeed well defined. That is it is independent of choice of parameterization

Theorem 179

Let $g : A \to B$ be a diffeomorphism of open sets in $\mathbb{R}^k$. Let $\beta : B \to \mathbb{R}^n$ be a map of class C^r. Let $Y = \beta(B)$. Let $\alpha = \beta \circ g$. then $\alpha : A \to \mathbb{R}^n$ and $Y = \alpha(A)$. If $f : Y \to \mathbb{R}$ is a continuous function, then f is integrable over Y_β if and onl if it is integrable over Y_α in this case

$$\int_{Y_\alpha} f\,dV = \int_{Y_\beta} f\,dV$$

where $v(Y_\alpha) = v(Y_\beta)$

Proof. By definition this is equivalent to showing that

$$\int_B (f \circ \beta)V(D\beta) = \int_A (f \circ \alpha)V(D\alpha)$$

The change of variables theorem tells us that

$$\int_B (f \circ \beta)V(D\beta) = \int_A (f \circ \beta \circ g)V(D\beta \circ g)\,|\det Dg|$$

so if just have to show that

$$V(D\beta \circ g)\,|\det Dg| = V(D\alpha)$$

to prove the theorem. Note that by chain rule 13

$$D\alpha(x) = D\beta(y)Dg(x)$$

Hence

$$[V(D\alpha(x))]^2 = \det(Dg(x)^t D\beta(y)^t D\beta(y)Dg(x)) = \det(Dg(x))^2[V(D\beta(y))]^2$$

as desired

Example 180

Let A be an open interval in $\mathbb{R}^1$ and let $a : A \to \mathbb{R}^n$ be a map of C^r. Let $Y = a(A)$. Then Y_a is called a parameterized-curve in $\mathbb{R}^n$. It's volume is 1D is what we call the **length** at it is given by the **line integral** below

$$v(Y_a) = \int_A V(Da) = \int_A \left[\left(\frac{da_1}{dt}\right)^2, \dots, \left(\frac{da_n}{dt}\right)^2\right]^{\frac{1}{2}} = \int_A ||Da||$$

For example consider the parameterized curve

$$a(t) = (a\cos t, a\sin t) \text{ for } 0 < t < 3\pi$$

where we have

$$Da = \left[\frac{\partial a}{\partial t}\right] = \left[-a\sin t, a\cos t\right]^t$$

so

$$\int_{Y_a} ds = \int_0^{3\pi} [a^2\sin^2 t + a^2\cos^2 t]^{\frac{1}{2}}$$

Example 181

Similar to the above example now let A be open in $\mathbb{R}^2$. Then the 2-dimensional volume is what we call the **area**. In the special case where $n = 3$ that is $a : A \to \mathbb{R}^3$ then we can write

$$v(Y_a) = \int_A V(Da) = \int_A \left\| \frac{\partial a}{\partial x} \times \frac{\partial a}{\partial y} \right\|$$

where the expression in the right most integral is literally the magnitude of the cross product. For example consider the graph

$$a(x, y) = (x, y, f(x, y))$$

so we have

$$Da = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}$$

you may verify that

$$\det\left[(Da)^t Da\right] = \left\| \begin{bmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x} \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y} \end{bmatrix} \right\| = \left\| \frac{\partial a}{\partial x} \times \frac{\partial a}{\partial y} \right\| = 1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2$$

so therefore

$$v(Y_a) = \int_A \left[1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right]^{\frac{1}{2}}$$

8 manifolds

Definition 182

Let $k < n$. The **canonical submersion map** π is defined by

$$\pi : \mathbb{R}^n \to \mathbb{R}^k, \quad (x_1, \ldots, x_n) \to (x_1, \ldots, x_k)$$

The **canonical immersion map** ι is defined by

$$\pi : \mathbb{R}^k \to \mathbb{R}^n, \quad (x_1, \ldots, x_k) \to (x_1, \ldots, x_k, 0, \ldots, 0)$$

8.1 immersion submersion theorems

Lemma 183 (Canonical Sumbermision lemma)

Let $A : \mathbb{R}^n \to \mathbb{R}^k$ be a linea map and suppose that A is onto. Then there exists a bijective linera map $B : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$A \circ B = \pi$$

Proof. Let $v_1, \ldots, v_n$ be a basis of $\mathbb{R}^n$ such that $Av_i = e_i, i = 1, \ldots, k$ is a basis of $\mathbb{R}^k$ and $Av_i = 0$ for all $i > k$. Now let $B : \mathbb{R}^n \to \mathbb{R}^n$ be the map $Be_i = v_i, i = 1, \ldots, n$ To visualize all these consider

$$A(v_1, \ldots, v_n) = (e_1, \ldots, e_k)$$

which is not necessarily **canonical**. Then there exists

$$A(B(e_1, \ldots, e_n)) = (e_1, \ldots, e_k)$$

So now $A \circ B = \pi$

Lemma 184 (Canonical Immerison lemma)

As before let $k < n$. Let $A : \mathbb{R}^k \to \mathbb{R}^n$ be a one-to-one linear map. Then there exists a bijective linera map $B : \mathbb{R}^n \to \mathbb{R}^n$ such that $B \circ A = \iota$.

Proof. Note that $B \circ A = \iota \Leftrightarrow A^t B^t = \pi$. Now use *canocial submersion lemma* above

Definition 185

Let $k < n$ The map f is a **submersion** at p if $Df(p) : \mathbb{R}^n \to \mathbb{R}^k$ is onto. A point $a \in \mathbb{R}^k$ is a regular value of f if for every point $p \in f^{-1}(a)$, f is a submersion at p.

Theorem 186 (Canonical Submersion Theorem)

Let U be an open set in $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^k$ be a C^∞ map. Let $p \in U$ and $k < n$.

Assume that f is a submersion at p and that $f(p) = 0$ then there exists a neighbourhood U_0 of p in U , a neighbourhood V of 0 in $\mathbb{R}^n$ and a diffeomorphism $g : V \to U_0$ such that $f \circ g = \pi$

First define $h : U \to \mathbb{R}^n$ by Since by assumption f is a submersion at p so replacing f with $f \circ T_p$ where $T_p : x \to x + p$ meaning that $Df(0) = A : \mathbb{R}^n \to \mathbb{R}^k$ is onto. So by **canonical submersion lemma** there exists $A \circ B$ where $B : \mathbb{R}^n \to \mathbb{R}^n$ is bijective map. That is

$$Df(B(0)) = D(f \circ B)(0) = \pi$$

Now you see why we replaced we used $Df(0)$ instead of directly $Df(p)$ where we will end up with $Df(p(0)) \to Df(B(p'))$. This is because 0 is canonical, no matter what basis vectors you use there is only 1 way to represent it, that is all their coefficients are 0. Now define

$$h(x_1, \ldots, x_n) = (f \circ B(x_1, \ldots, x_k), x_{k+1}, \ldots, x_n)) = (x_1, \ldots x_n)$$

$$Dh(x_1, \ldots, x_n) = \begin{bmatrix} D_1 h, \ldots D_n h \end{bmatrix} \mathbf{x} = (D(f \circ B)(x_1, \ldots, x_k), D(x_{k+1}, \ldots, x_n)) = D(x_1, \ldots x_n)$$

So $Dh(0) = I$. To see this consider

$$\begin{bmatrix} \frac{\partial h_i}{\partial x_j}(0) & \cdots \\ \vdots & \ddots \\ & \end{bmatrix} = \begin{bmatrix} \frac{\partial x_i}{\partial x_j}(0) & \cdots \\ \vdots & \ddots \\ & \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \\ 0 & & 1 \end{bmatrix}$$

Recall linear algebra taught in `Artin`, since there exists a basis where the matrix of transformation is the identity matrix this is a bijective map. Hence by **inverse function theorem**, the function h maps a neighbourhood of U_0 of 0 diffeomorphically onto a neighbourhood V of 0 in $\mathbb{R}^n$. Therefore we know

$$f \circ h^{-1} = \pi$$

where $g = h^{-1}$.

Definition 187
Let $k < n$ The map f is a **immersion** at p if $Df(p) : \mathbb{R}^k \to \mathbb{R}^n$ is one to one

Theorem 188 (Canononical Immersion Theorem)
Let U be a neighbourhood of 0 in $\mathbb{R}^k$ and let $f : U \to \mathbb{R}^n$ be a C^∞ map. Assume that f is an immersion at 0. Then there exists a neighbourhood of $f(0) = p$ in $\mathbb{R}^n$ a neighbourhood W of 0 in $\mathbb{R}^k$ and a diffeomorphism $g : V \to W$ such that $\iota^{-1}(W) \subseteq U$ and $g \circ f = \iota$

Proof. Again just take the transpose matrix maps and use canonical submersion theorem.

Definition 189
A subset $X \subseteq \mathbb{R}^N$ is an **n-dimensional manifold** if for every point $p \in X$ there exists a neighbourhood of V of $p \in \mathbb{R}^N$, an open set $U \in \mathbb{R}^n$ and a diffeomorphism $f : U \to V \cap X$. The collection (f, U, X) is called a **parameterization** of X at p

Fact 190
Parametrizations are **smooth**, since recall the map $f : X \to Y$ is a diffeomorphism if it is bijective and both f and f^{-1} are $\mathcal{C}^\infty$ maps

Definition 191 (Smooth map)
Let S be a subset of $\mathbb{R}^k$ let $f : S \to \mathbb{R}^n$. We say that f is of class C^r on S if f may be extended to a function $g : U \to \mathbb{R}^n$ that is of class C^r on an open U of $\mathbb{R}^k$ containing S.

Proof. Cover S by the open neighbourhoods $\{U_x\}$. Let A be its union. Then let $\{\phi_i\}$ be a partition of unity subordinate to $\{U_x\}$. Define a C^r function $g_i : U_x \to \mathbb{R}^n$ that has the same support as ϕ_i. So we ave $\phi_i g_i : U_x \to \mathbb{R}^n$. Because by definition we have $\operatorname{supp} \phi_i g_i \subseteq U_x$, we extend these functions to vanish outside U_x instead(recall we like how we modify the intervals in our bump functions 2.4). In other words we want the support($g \neq 0$) of the function to extended to all of A not just a subset of A. The reason for doing so is to ensure with certainty that were f is supported on S, so does g.So we don't end up with situation where $\phi_i(x) = 0, \forall i$ but $f(x) \neq 0$ for the same $x \in S$. So we define h_i to be such an extension. Therefore we can now define

$$g(x) = \sum_{i=1}^{\infty} h_i(x)$$

for $x \in A$. It is clear $g(x) \in C^r$. Now suppose that $f(x)$ agrees with f on $U_x \cap S$. Then

$$h_i(x) = \phi_i(x) g_i(x) = \phi(x)_i f(x)$$

for $x \in S$. Therefore we can say

$$g(x) = \sum_{i=1}^{\infty} \phi_i(x) f(x) = f(x)$$

for $x \in S$ where can conclude $f \in C^r$. To be precise our extension is function is defined as

$$g(x) = \begin{cases} g(x) & x \in A/S \\ f(x) & x \in S \end{cases}$$

Remark 192. *Earlier we have defined $\int_A f = \sum_i \int_A \phi_i f$ assuming the RHS converges. We just dealt with a new case, that is $f(x) = \sum \phi_i f(x)$*

Lemma 193 (Smoothness)
Let X be a subset of $\mathbb{R}^m$ and $f : X \to \mathbb{R}^n$ The map f is $\mathcal{C}^r$ if for every $p \in X$ there a neighbourhood U_p of p in $\mathbb{R}^m$ and a $\mathcal{C}^r$ map $g_p : U_p \to \mathbb{R}^n$ such that $g_p = f$ on $U_p \cap X$

Theorem 194 (Inverse Images of regular values are manifolds)
Let U be an open subset of $\mathbb{R}^N$ let $k < N$ and let $f : \mathbb{R}^N \to \mathbb{R}^k$ be a $\mathcal{C}^\infty$ map. Suppose that 0 is a regular value of f that implies if you recall

$$f^{-1}(0) \cap C_f = \emptyset$$

Then, the set $X = f^{-1}(0)$ is an n-dimensional manifold where $n = N - k$

Remark 195. *Note for f^{-1} and π^{-1}, we are referring to their pre-images. This is not proper writing conventionally because the inverse is only defined if bijective. Just bear this in mind for this case.*

Proof. If $p \in f^{-1}(0) = X$ then $p \notin C_f$. So the map $Df(p) : \mathbb{R}^N \to \mathbb{R}^k$ is onto so it is a submersion at p. By canoncial submersion theorem there exists a neighbourhood V of 0 in $\mathbb{R}^N$ a neighbourhood U_0 of p in U and a diffeomorphism $g : V \to U_0 \in \mathbb{R}^N$ such that

$$f \circ g = \pi$$

where $\pi : \mathbb{R}^N \to \mathbb{R}^k$ Consider the pre-image of π of $0 \in \mathbb{R}^k$. That is $\pi^{-1} = \{x : \pi(x) = 0\}$.

$$\pi^{-1}(0) = g^{-1} f^{-1}(0) = g^{-1}(X) = \{0\} \times \mathbb{R}^n = \mathbb{R}^n$$

because g is bijective we infer that X has dimension n. hence combining $g : \pi^{-1}(0) \to f^{-1}(0)$ with $g : V \to U_0$ we have

$$g \text{ maps } V \cap \pi^{-1}(0) \text{ diffeomorphically onto } U_0 \cap X$$

where U_0 is a neighbourhood of p. Moreover $V \cap \pi^{-1}(0)$ and $U_0 \cap X$ are clearly an open sets in $\mathbb{R}^n$ □

We now consider consider examples of applications of theorem 194

Example 196
We consider the n-sphere S^n. Define a map

$$f:\mathbb{R}^{n+1}\to\mathbb{R},\quad f(x)=x_1^2+\dots x_{n+1}^2-1$$

The derivative is $(Df)(x)=2[x_1,\dots,x_{n+1}]$ so $C_f=\{0\}$ as $\det D(f)(0)=0$ it is clearly not onto so it cant be bijective. To make use of the previous suppose $a\in f^{-1}(0)$ then $\sum a_i^2=1$ so it is not in C_f. Thus having proved that a is a regular value the previous theorem implies that $f^{-1}(0)=S^n$ is an n-dimensional manifolds

Example 197 (Graphs are manifolds)
Let $g:\mathbb{R}^n\to\mathbb{R}^k$ be a $\mathcal{C}^\infty$ map. Define

$$X=\text{graph } g=\left\{(x,y)\in\mathbb{R}^n\times\mathbb{R}^k : y=g(x)\right\}$$

Note that $X\subseteq\mathbb{R}^n\times\mathbb{R}^k=\mathbb{R}^{n+k}$ The set X is an n-dimensional manifold.

Proof. Define the map $f:\mathbb{R}^n\times\mathbb{R}^k\to\mathbb{R}^k$ b7

$$f(x,y)=y-g(x)=0$$

Then we see that $Df(x,y)=[-Dg(x),I_k]$.By the second term I_k we realize that the matrix $Df(x,y)$ has rank k so it is surjective/onto so 0 is a regular point. Hence the graph $g=f^{-1}(0)$ is an n-dimensional manifold of $\mathbb{R}^{n+k}$

Fact 198
Suppose $f:\mathbb{R}^N\to\mathbb{R}^k$ is a continuous map and suppose 0(so $\deg f\neq 0$ recall 151) is a regular value so $X=f^{-1}(0)$ is an n-dimensional manifold. Suppose that $f=(f_1,\dots,f_k)$ then X can be defined by

$$f_i(x_1,\dots,x_N)=0,\quad i=1,\dots,k$$

Which are a system of **non-degenerate**(injective) equations since for every $x\in X$ the matrix

$$\left[\frac{\partial f_i}{\partial x_i}(x)\right]$$

is of rank k(surjective,full column rank). We will prove that every n-dimensional manifold $X\subseteq\mathbb{R}^N$ can be described locally by a system of k non-degenerate equations like so with the same above conditions

Lemma 199
Let $X\subseteq\mathbb{R}^N$ be an n-dimensional manifold, p a point of X and U open subset of $\mathbb{R}^n$, V a neighbourhood of p in $\mathbb{R}^N$ and the parameterization be $\varphi:U\to V\cap X$ be diffeomorphism. Modifying φ by translations if necessary we can assume that $0\in U$ and $\phi(0)=p$. The linear map $(D\varphi)(0):\mathbb{R}^n\to\mathbb{R}^N$ is injective - that is φ is an immersion at 0

Proof. $\varphi^{-1}:V\cap X\to U$ is a diffeomorphism so by definition of $\mathcal{C}^\infty$ functions there exists an extension $\psi:V\to U$ that coincides with φ^{-1} on $V\cap X$. Therefore by chain rule

$$D(\psi\circ\varphi)(0)=(D\psi)(p)D\varphi(0)=I$$

since $\varphi \circ \varphi^{-1} = \varphi \circ \psi$. Hence the form above implies there exists an inverse $(D\phi)(p)$. Equivalently if we write $D\phi(0)v = 0$ then $v = 0$ is the only solution to the nullspace here since $D\varphi(0)$ cant be zero as seen.

Theorem 200

et $X \subseteq \mathbb{R}^N$ be an n-dimensional manifold.For every $p \in X$ there exists a neighbourhood V_p of p in $\mathbb{R}^N$ an a submersion $f : V_p \to \mathbb{R}^k$ with $f(p) = 0$ such that X is defined locally near p by the equation

$$V_p \cap X = f^{-1}(0)$$

Proof. from 199 where $U_0 \in \mathbb{R}^n$ there we know that $\phi : U_0 \to V_p \cap X$ is an immersion at 0. Hence by canonical immerions theorem there exists a neighbourhood U_0 of 0 in U, a neighbourhood V_p of p in V a neighbohood $\mathcal{O}$ of 0 in $\mathbb{R}^N$ a diffeomorphism

$$g : (V_p, p) \to (\mathcal{O}, 0)$$

such that

$$\iota(U_0) = \mathcal{O}$$

and

$$g \circ \varphi = \iota$$

on U_0. So we have

$$\iota(U_0) = g \circ \varphi(U_0) = g(V_p \cap X) \subset \mathcal{O}$$

that is to say if $g = (g_1, \ldots, g_N)$ every solution $V_p \cap X$ satisfies

$$g_i = 0, \quad i = n+1, \ldots, N$$

Let $k = N - n$ and let

$$\pi : \mathbb{R}^N = \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^k$$

be the canonical submersion

$$\pi(x_1, \ldots, x_N) = (x_{n+1}, \ldots, x_N)$$

so let $f = \pi \circ g$. This makes f a submersion as g is a diffeomorphism on V_p which implies $g \in \mathcal{C}^\infty$ so $Dg(V_p) = [D_1 g(V_p), \ldots, D_N g(V_p)]$ exists. So $\pi(Dg(V_p)) = [D_{n+1} g(V_p), \ldots, D_N g(V_p)]$ is clearly onto. From the above therefore we have

$$f(V_p \cap X) = \pi \circ g(V_p \cap X) = 0$$

hence

$$V_p \cap X = f^{-1}(0)$$

8.2 tangent spaces of manifolds

Proposition 201

Let $X \subseteq \mathbb{R}^N$ be an n-dimensional manifold, p a point of X,U an open subset of $\mathbb{R}^N$, V a neighbourhood of p in $\mathbb{R}^N$. Again modifying φ by a translation if necessary such that $q \in U$ and $\phi^{-1}(p) = q$, define the parameterization be $\varphi : (U, q) \to (V \cap X, p)$ that is a diffeomorphism. Then the map $(d\varphi)_q : T_q\mathbb{R}^n \to T_p\mathbb{R}^N$ is one to one

Proof. Again by definition of $\mathcal{C}^\infty$ there exists $\psi = \varphi^{-1}$ where $\varphi : V \to U$ and conincides with φ^{-1} on $V \cap X$. Again by chain rule

$$d(\psi \circ \varphi)_q = (d\psi)_p \circ (d\varphi)_q = d(\mathrm{id}_U)_q$$

so $(d\varphi)_q$ is injective. Note here we write $\varphi : (U, q) \to (V \cap X, p)$ because we are only interested in the root of the 2 element tuple pairs in the tangent space.

Definition 202

Define $p \in X$ and $T_pX \subseteq T_p\mathbb{R}^N$ Let $\phi : U \to V \cap X$ be a parametirzation of X and let $\phi(q) = p$. We define the **tangent space** of a manifold X to be

$$T_pX = \mathrm{Im}(d\phi)_q$$

Recall that since $(d\phi)_q$ is injective, that means T_pX is n-dimensional by rank nullity theorem

Definition 203

An alternate defintition for the **tangent space** of a manifold is

$$T_pX = \ker df_p$$

where f is a function satisfying 200. These two definitions are equivalant

Proof. We know from 200 that $f \circ \phi = 0$ because of the sets that define their domain and codomains. Therefore

$$df_p \circ d\phi_q = d(f \circ \phi)_q = 0$$

So it is clear that $\mathrm{Im}\, d\phi_q = \ker df_p$

Remark 204. *We have shown that the definition of the tangent space is* independent *of the choice of parameterization because no matter what ϕ we define, both definitions will still hold.*

Lemma 205

Let W be an open subset of $\mathbb{R}^\ell$ and let $g : W \to \mathbb{R}^n$ be a $\mathcal{C}^\infty$ map. Suppose that $g(W) \subseteq X$ and that $g(w) = p$ where $w \in W$. Then $(dg)_w \subseteq T_pX$

Proof. We make use of 203 that there exists a sumbermersion $f : V \to \mathbb{R}^k$ where $X \cap V = f^{-1}(0)$. We express the $\ker df_p = T_pX$ by trying to end up with the set $X \cap V$. We can then define $W_1 = g^{-1}(V)$. Since $g(W) \subseteq X$ where W is the whole domain of g as defined, then $g(W_1) \subseteq V \cap X$ must clearly be satisfied. Therefore we can now let $w \in W_1$ and we know $df_p \circ dg_w = 0$. Therefore $dg_w \subseteq T_pX$ since dg_w is the nullspace of df_p.

Fact 206

Suppose now $X \subseteq \mathbb{R}^N$ is an n-dimensional manifold and $Y \subseteq \mathbb{R}^\ell$ is an m-dimensional manifold. Let $f : X \to Y$ be a $\mathcal{C}^\infty$ map. Let $q = f(p)$. So we define the map

$$df_p : T_pX \to T_qY$$

This makes sense if we let V be neighbourhood of p in $\mathbb{R}^N$ and let $g : V \to \mathbb{R}^\ell$ be map such that $g = f$ on $V \cap X$. By definition that means $T_pX \subseteq T_p\mathbb{R}^N$ so when we have

$$dg_p : T_p\mathbb{R}^N \to T_q\mathbb{R}^k$$

we define df_p to simply be the restirction of dg_p to the tagent space T_pX. Naturally you might consider the following for this to be well-defined definition

Proposition 207

The following properties hold

1. Im $dg_p(T_pX)$ is a subset of T_qY
2. The definition does not depend on the choice of g
3. (Chain Rule Manifolds) $d(g \circ f)_p = (dg_q) \circ (df)_q$

Proof. see notes and chatgpt to be continued...

Example 208

Consider the following

$$\mu \text{ is a k-form on } X$$
$$\nu \text{ is an } \ell \text{ form on } X.$$

For $p \in X$ we have

$$\mu_p \in \Lambda^k(T_p^*X)$$
$$\nu_p \in \Lambda^\ell(T_p^*X)$$

Taking the wedge product we have

$$\mu_p \wedge \nu_p \in \Lambda^{k+\ell}(T_p^*X)$$

From this we get the pullback

$$(df_p)^* : \Lambda^k(T_q^*Y) \to \Lambda^k(T_p^*X)$$

then f^*w is defined by

$$(f^*w)_p = (df_p)^* w_q$$

They are all related by

$$\begin{tikzcd} T_pX \arrow[r, "df_p"] & T_qY \\ T_0U_1 \arrow[u, "d(\phi_1)_0"] \arrow[r, "d\psi_0"] & T_0U_2 \arrow[u, "(d\phi_2)_0"'] \end{tikzcd}$$

Definition 209

The space of $\mathcal{C}^\infty$ k-forms on X is denoted by $\Omega^k(X)$ Let $w \in \Omega^k(X)$. Let $X \subseteq \mathbb{R}^m$ be an n-dimensional manifold.Let U is open subset of $\mathbb{R}^N$ and V is a neighbourhood of p in $\mathbb{R}^N$ and the diffeomorphism $\phi : U \to V \cap X$ be the parameterization of X at p.

1. The k form w is $\mathcal{C}^\infty$ at p if there exists a k-form $\tilde{w} \in \Omega^k(V)$ such that $\iota_X^* \tilde{w} = w$
2. The k form w is $\mathcal{C}^\infty$ at p if there exists a diffeomorphism $\phi : U \to V \cap X$ such that $\phi^* w \in \Omega^k(U)$

The above 2 definitions of $\mathcal{C}^\infty$ k-forms w on manifolds are equivalant.

Proof. First we know that ι_X is C^∞ since for every $p \in X$ the indentity map $I : \mathbb{R}^n \to \mathbb{R}^n$ which is clearly C^∞ coincides with ι_X on X. Now, the first definition implies the second definition if you consider

$$\phi^* w = \phi^*(\iota_X^* \tilde{w}) = (\iota_X \circ \phi)^* \tilde{w}$$

by chain rule. Clearly the map $\iota \circ \phi : U \to V$ is C^∞ too and $\tilde{w} \in \Omega^k(V)$. Therefore

$$\phi^* w = (\iota_X \circ \phi)^* \tilde{w} \in \Omega^k(U)$$

. To prove the other direction that the second definition implies the first again. Let $\phi : U \to V \cap X$ be a diffeomorhism. Then we know as usual $\phi^{-1} : V \cap X \to U$ can be extended to $\psi : V \to U$ where ϕ^{-1} conincides with ψ on $V \cap X$. Consider the k-form $\tilde{w}$ on $V \subset \mathbb{R}^N$ defined by $\tilde{w} = \psi^*(w)$ which implies $\tilde{w}$ is a smooth k-form in $\Omega^k(V)$. Now, consider the inclusion map $\iota_X : X \to \mathbb{R}^N$. By construction, $\tilde{w}$ pulls back under ι_X^* to give w. This is because on U, $\phi^* \tilde{w} = \phi^*(\psi^* w) = w$.

Remark 210. *Once again f is independent choice of pasteurization both definitions hold regardless of it*

Theorem 211

If $w \in \Omega^k(X)$ then there exsts a neighbourhood W of X in $\mathbb{R}^N$ and a k-form $\tilde{w} \in \Omega^k(W)$ such that $\iota_X^* \tilde{w} = w$

Proof. Let $p \in X$. There exists a neighbourhood V_p of p in $\mathbb{R}^N$ and a k-form $w^p \in \Omega^k(V_p)$ such that $\iota_X^* w^p = w$ on $V_p \cap X$. Let

$$W \subseteq \bigcup_{p \in X} V_p$$

THe collection of sets $\{V_p : p \in X\}$ is an open cover of w Let $p_i, i = 1, 2, 3 \ldots$ be a partition of unity subordinate to this cover. So $p_i \in \mathcal{C}_0^\infty$ and $\operatorname{supp} p_i \subset V_p$ for some p let

$$\tilde{w}_i = \begin{cases} p_i w^p & \text{on } V_p \\ 0 & \text{elsewhere} \end{cases}$$

Notice upon restriction from W to X

$$\iota_X^* \tilde{w}_i = \iota_X^* p_i \iota_X^* w^p = (\iota_X^* p_i) w$$

where recall this is how you distribute the pullback operators. Take

$$\tilde{w} = \sum_{i=1}^{\infty} \tilde{w}_i = \sum_{i=1}^{M} p_i w$$

for some finite M.This makes sense since we used a partition of unity among a zero form(p_i) and non-zero form. From th sum we can see that $\tilde{w} \in \Omega^k(W)$ finally

$$\iota_X^* \tilde{w} = \left(\iota_X^* \sum p_i\right) w = w$$

because $\iota_X^* w = w$ since $w \in \Omega^k(X)$ already

Theorem 212

If $\mu, \nu \in \Omega^k(Y)$ then

$$f^*(\mu \wedge \nu) = f^*\mu \wedge f^*\nu$$

and differential operation d on k-form manifolds can be defined by

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

Proof. Having properly defined tangent spaces, and what it means to be C^∞ k-forms for manifolds,chain rule, you should have realized all other properties are exactly the same as euclidean

8.3 Orientation on Manifolds

Definition 213

Suppose $\mathbb{L}$ be a one-dimensional vector space and so the set $\mathbb{L} - \{0\}$ has the 2 components

$$\mathbb{L}_+ = \{\lambda v : \lambda > 0\} \text{ and } \mathbb{L}_- = \{\lambda v : \lambda < 0\}$$

which correspond to the *choice of orientation*

Definition 214

An **orientation** of V is an orientation of the one-dimensional vector space $\Lambda^n(V^*)$. That is, an orientation of V is a choice of $\lambda^n(V^*)_+$

Definition 215

Suppose V_1, V_2 are oriented n-dimensional vector spaces and let $A : V_1 \to V_2$ be a bijective map. Then the map A is **orientation preserving** if

$$w \in \Lambda^n(V_2)_+ \Rightarrow A^* w \in \Lambda^n(V_1)_+$$

We now aim to generalize the notion of oreienttation to manifolds. Suppose $X \subset \mathbb{R}^N$ be an n-dimensinoal manifold.

Definition 216

An orientation of X is a function on X which assigns to each point $p \in X$ an orientation of T_pX

Example 217

Let $w \in \Lambda^n(X)$ and suppose that w is nowhere vanishing. Orient X by assigning t $p \in X$ the orientation of T_pX for which $w_p \in \Lambda^n(T_p^*X)_+$.

Example 218

Let U be an open subset of $\mathbb{R}^n$. The **standard orientation** of U is the canonical orientation defined by the euclidean volume element $dx_1 \wedge \ldots \wedge dx_n$

Definition 219 ($\mathcal{C}^\infty$ orientations)

An orientation of X is a $\mathcal{C}^\infty$ orientation if for every point $p \in X$ there exists a neighbourhood U of p in X and an n-form $w \in \Omega^n(U)$ such that all points $q \in U, w_q \in \Lambda^n(T_q^*X)_+$

From now on we only consider $\mathcal{C}^\infty$ orientations.

Theorem 220 (Oriented Manifolds)

If X is **oriented** then there exists $w \in \Omega^n(X)$ such that for all $p \in X, w_p \in \Lambda^n(T_p^*X)_+$

Proof. For every point $p \in X$ there exists a neighbourhood U_p of p and an n-form $w^p \in \Omega^n(U_p)$ such that for all $q \in U_p, (w^p)_q \in \Lambda^n(T_q^*X)_+$. Take $p_i, i = 1, 2, \ldots$ to be a partiion of unity suborindate to $\mathbf{u} = \{U_p : p \in X\}$. FOr every i such that $p_i \in \mathcal{C}_0^\infty(U_p)$, let

$$w_i = \begin{cases} p_i w^p & \text{on } U_p \\ 0 & \text{on } X - U_p \end{cases}$$

Since both w^p and p_i are smooth, w_i is a $\mathcal{C}^\infty$ map.We have thus constructed a w that is positively oriented a ever point

Remark 221. *Such an n-form $w \in \Omega^n(X)$ with this property are known as* **volume forms**

Definition 222 (Orientation Preserving Diffeomorphisms)

Let X, Y be oriented manifolds. A diffeomorphism $f : X \to Y$ is **orientation preserving** if for every $p \in X$

$$(df_p)^* : \Lambda^n(T_q^*)_+ \to \Lambda^n(T_p^*)_+$$

is orientation preserving where $q = f(p)$

Definition 223 (Orientation Preserving Parameterizations)

Let X be an oriented manifold. Let U be an open subset of $\mathbb{R}^n$ and V be an open subset of X. Let the $\phi : U \to V$ be the parameterization which is a diffeomorphism. ϕ is an **oriented parameterization** if it is orientation preserving.

Note that if X is an oriented manifold. Then there exists for every $p \in X$ an oriented parameterization of X at p since even if ϕ was orientation reversing we can always replace ϕ by $\phi'(x_1, \ldots, x_n) = \phi(x_1, \ldots, -x_n)$ which will reverse the orientation back.

8.4 Integration on Manifolds

Definition 224

Let $\phi : U \to V$ be an oriented parameterization. Let U be open in $\mathbb{R}^n$ and let V be open in X. Take any $w \in \Omega_c^n(V)$. Then

$$\int_V w = \int_U \phi^* w$$

where $\phi^* w = f(x) dx_1 \wedge \ldots \wedge dx_n$ where $f \in \mathcal{C}_0^\infty(U)$ and

$$\int_U \phi^* w = \int_U f$$

This definition of $\int w$ is independent of the choice of oriented parameterization ϕ

Proof. Let $\phi_i : U_i \to V_i$, $i = 1, 2$ be oriented paramterizations. Let $w \in \Omega_c^n(V_1 \cap V_2)$. Define

$$U_{1,2} = \phi_1^{-1}(V_1 \cap V_2)$$
$$U_{2,1} = \phi_2^{-1}(V_1 \cap V_2)$$

Since both ϕ_1 and π_2 are diffeomorphisms we have the diagram

$$\begin{tikzcd} V_1 \cap V_2 \arrow[r, equal] & V_1 \cap V_2 \\ U_{1,2} \arrow[u, "\phi_1"] \arrow[r, "f"] & U_{2,1} \arrow[u, "\phi_2"'] \end{tikzcd}$$

Therefore $f = \phi_2^{-1} \circ \phi_1$ is a diffeomorphism and $\phi_1 = \phi_2 \circ f$. Integratig we have

$$\begin{aligned} \int_{U_1} \phi_1^* w &= \int_{U_{1,2}} \phi_1^* w \\ &= \int_{U_{1,2}} (\phi_2 \circ f)^* w \\ &= \int_{U_{1,2}} f^*(\phi_2^* w) \end{aligned}$$

But because ϕ_1 and ϕ_2 are orientation preserving, so is f therefore Therefore we have

$$\int_{V_1} w = \int_{U_{1,2}} \phi_1^* w = \int_{U_{2,1}} \phi_2^* w = \int_{V_2} w$$

$$\int_{U_{1,2}} f^* \phi_2^* w = \int_{U_{2,1}} \phi_2^* w$$

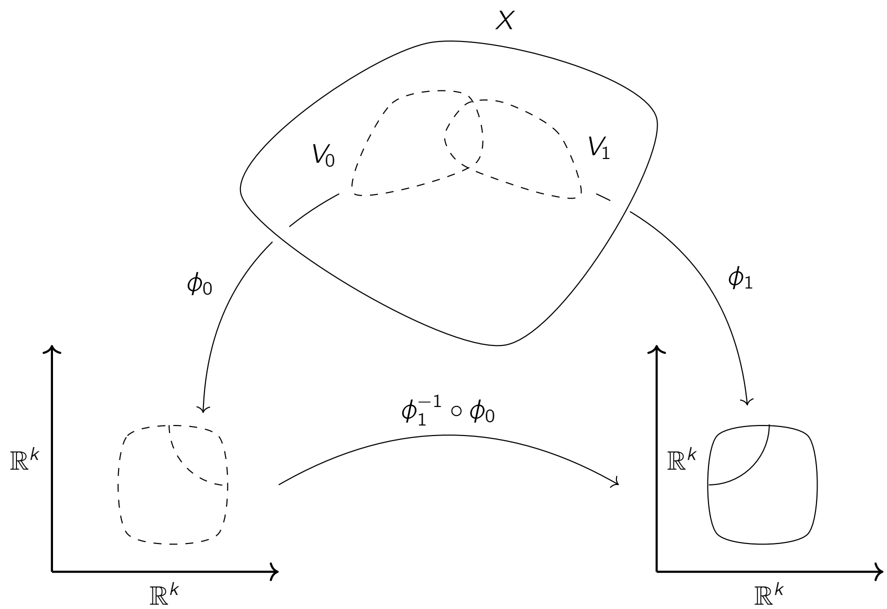


Figure 1: **transition functions** on manifolds

Definition 225

Let $\{V_\alpha : \alpha \in U\}$ be an open covering of X by parametizable open sets where $p_i \in \mathcal{C}^\infty(X)$ is a partiion of unity subordinate to V_α. Then for any compactly supported n-form $w \in \Omega_c^n(X)$ we may define the integral

$$\int_X w = \sum_{i=1}^{\infty} \int_{V_i} p_i w$$

This definition is independent of the choice of partition of unity.

Proof. Let $\{U'_\alpha : \alpha \in U'\}$ be a diferent open covering of X by parameterizable open sets and p'_i be the corresponding partition of unity then by swapping the order of summation we have

$$\sum_i \int_X p_i w = \sum_{i,j} \int_X p_i p'_j w = \sum_{j,i} \int_X p'_j p_i w = \sum_j \int_X p'_j w$$

Theorem 226 (Poincare Lemma Manifolds)

Let $X \subseteq \mathbb{R}^N$ be an *oriented connected* n-dimensional manifold. Then For any $w \in \Omega_c^n(X)$ the following are equivalent

1. $\int_X w = 0$
2. $w \in d\Omega_c^{n-1}(X)$

We will only prove (1) implies (2). The proof for (2) implying (1) is a consequence of **stokes theorem** which we shall learn later. First consider the following 2 lemmas.

Lemma 227 (Connectivity Lemma)

Given $p, q \in X$ there exists open sets $W_j, j = 0, \ldots, N+1$ such that W_j is diffeomorphic to an open set in $\mathbb{R}^n$ and such that $p \in W_0$, $q \in w_{n+1}$ and that $W_i \cap W_{i+1} \neq \emptyset$

Proof. Fix a point $p \in X$. Consider the set S_p defined as:

$$S_p = \{q \in X \mid \text{there exist open sets } W_j, j = 0, \ldots, N+1, \text{ with the required properties}\}.$$

We need to show that $S_p = X$. Consider $q \in S_p$. By the definition of S_p, there exist open sets W_j, $j = 0, \ldots, N+1$, such that W_0 contains p, W_{N+1} contains q, each W_j is diffeomorphic to an open set in $\mathbb{R}^n$, and $W_j \cap W_{j+1} \neq \emptyset$ for all $j = 0, \ldots, N$.

Since W_{N+1} is an open set containing q, there exists an open neighborhood $U \subseteq W_{N+1}$ containing q. For any $q' \in U$, we can extend the sequence of open sets W_j by replacing W_{N+1} with U. Thus, $q' \in S_p$, which shows that S_p is open.

Now consider the complement $S_p^c = X \setminus S_p$. Assume $q \in S_p^c$. By the definition of S_p^c, there do not exist open sets W_j satisfying the conditions with W_0 containing p and W_{N+1} containing q.

Since X is a manifold, q has a neighborhood U diffeomorphic to an open set in $\mathbb{R}^n$. If any point $q' \in U$ were in S_p, there would exist a sequence of open sets connecting p to q', which could be extended to connect p to q by including U. This would imply $q \in S_p$, contradicting our assumption that $q \in S_p^c$. Therefore, $U \subseteq S_p^c$, and S_p^c is open.

We have shown that both S_p and S_p^c are open sets in X. Since X is connected and $S_p \cup S_p^c = X$, either $S_p = X$ or $S_p^c = X$.

By definition, $p \in S_p$, so S_p is non-empty. Therefore, S_p^c must be empty, which implies $S_p = X$. $\square$

Lemma 228

Like for eucildean we denote the equivalence relation for $w_1, w_2 \in \Omega_c^n(X)$

$$\int_X w_1 = \int_X w_2 \quad \equiv \quad w_1 \sim w_2 \quad \Leftrightarrow \quad w_1 - w_2 \in d\Omega_c^{n-1}(X)$$

The theorem is true if $V = V'$

Proof. Let $\phi : U \to V$ be an orientation preseving paramterization. If $w_1 \sim w_2$ then

$$\int \phi^* w_1 = \int \phi^* w_2$$

which is the same as saying that

$$\phi^* w_1 - \phi^* w_2 \in d\Omega_C^{n-1}(U)$$

therefore

$$w_1 - w_2 \in d\Omega_C^{n-1}(U)$$

Now back to poincare lemma for manifolds.

8.5 Integration on smooth domains

Example 229
The prototypal smooth domain(manifold with boundary) is the half plane

$$\mathbb{H}^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 \leq 0\}$$

and the boundary of the half plane is

$$\mathrm{Bd}\,\mathbb{H}^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 = 0\}$$

and the interior of the half plane is

$$\mathbb{H}^n_+ = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 < 0\}$$

Definition 230
Let X be an n-dimensional manifold. A closed subset $D \subseteq X$ is a **smooth domain** if for every point $p \in \mathrm{Bd}(D)$ there exists a parameterization $\phi : U \to V$ on X at p such that $\phi : U \cap \mathbb{H}^n \to V \cap D$ is a diffeomorphism. In other words ϕ is a *parameterization of D at p* too

Proof. Note that $\phi : U \cap \mathbb{H}^n \to V \cap D$ is a homeomorphism meaning it maps boundary points to boundary points. $\phi : U \cap \mathrm{Bd}\,\mathbb{H}^n \to V \cap \mathrm{Bd}\,D$

Proposition 231
The set $D = \{x \in X : f(x) \leq 0\}$ is a *smooth* domain.

Proof. Take $p \in \mathrm{Bd}(D)$ so $p = f^{-1}(0)$. Let $\phi : U \to V$ be a parameterization of X at p

8.6 Boundary of Manifold

Definition 232
A **k-manifold** in $\mathbb{R}^n$ of class C^r is a subspace of M of $\mathbb{R}^n$ having the following property. For each $p \in M$ there is an open set V of M containing p a set U that is open in either $\mathbb{R}^k$ or $\mathbb{H}^k$ and a continuous map $a : U \to V$ carryinng U onto V in a one-to-one fashion such that

1. a is of class C^r
2. $a^{-1} : V \to U$ is continuous
3. $Da(x)$ has rank k for each $x \in U$

The map a is called a **coordinate patch** on M about p

The zero manifold is basically just a set of discrete points in $\mathbb{R}^n$ while a manifold without boundary is simply the special case where all coordinate patches have domains open in $\mathbb{R}^k$

Example 233

The set $\mathbb{R}^{k-1} \times 0$ is open in $\mathbb{H}^k$ but not open in $\mathbb{R}^k$. Just consider the points at the boundary of $\mathbb{R}^{k-1} \times 0$, i.e when $x_i = \mathbb{R}, i = 1, \ldots, k$ and $x_k = 0$. To define a neighbourhood we must have

$$(\mathbb{R} \pm \varepsilon_1, \ldots, \mathbb{R} \pm \varepsilon_{k-1}, 0 \pm \varepsilon_k)$$

But in $\mathbb{R}^k$ this set here is no longer contained in $\mathbb{R}^{k-1} \times 0$. Hence there is no neighbour or radius $r > 0$ around those points in $\mathbb{R}^k$ that is contained in $\mathbb{R}^{k-1} \times 0$!

Refer back to `rudin` if necessary, it is very same idea why the metric space X itself is both open and closed sets of itself.

Definition 234

Let M be a k-manifold in $\mathbb{R}^n$ let $p \in M$. If there is a coordinate patch $a : U \to V$ on M about p such that U is open in $\mathbb{R}^k$ we say p is an **interior point** of M.Otherwise we say p is a **boundary point** of M. We denote the set of boundary points of M by ∂M

Lemma 235

Let M be a k-manifold in $\mathbb{R}^n$ let $a : U \to V$ be a coordinate patch about the point p of M.

(a) If U is open in $\mathbb{R}^k$ then p is an interior point of M

(b) If U is open in $\mathbb{H}^k$ and if $p = a(x_0)$ for $x_0 \in \mathbb{H}^k_+$ thenn p is an interir point of M

(c) If U is open in $\mathbb{H}^k$ and $p = a(x_0)$ for $x_0 \in \mathbb{R}^{k-1} \times 0$ then p is a boundary point of M

Proof. (a) is immediate from dfintion. For (b) given $a : U \to V$, let $U_0 = U \cap H^k_+$ and let V_0 be $a(U_0)$. Clearly the mapping by $a|U_0$ of U_0 onto V_0 is a coordinate patch about p with U_0 open in $\mathbb{R}^k$. For (c) consider $a_0 : U_0 \to V_0$ be a coordinate path about p with U_0 is open in $\mathbb{H}^k$ and $p = a_0(x_0)$ for $x_0 \in \mathbb{R}^{k-1} \times 0$. Let the coordinate patch $a_1 : U_1 \to V_1$ about p with U_1 open in $\mathbb{R}^k$. $W = V_0 \cap V_1$ non empty(contains p). Let $W_i = a_i^{-1}(W)$. So we know W_0 is open in $\mathbb{R}^k$ while W_1 is open in $\mathbb{H}^k$ as they are contained in U_1 and U_0 respectively and W is an open set and thus so must their image sets given a_1, a_0 are C^r maps. Also the map

$$a_1^{-1} \circ a_0 : W_0 \to W_1$$

is non-singular as shown in 224 and clearly also C^r(composition of C^r functions). So essentially

$$\begin{tikzcd}
V_0 \cap V_1 \arrow[r, equal] & V_1 \cap V_0 \\
W_0 \arrow[u, "a_1"] \arrow[r, "a_1^{-1} \circ a_0"] & W_1 \arrow[u, "a_2"']
\end{tikzcd}$$

is what is going on. However x_0 is of $\mathbb{R}^{k-1} \times 0$ it is not open in $\mathbb{R}^k$ but this is a contradiction because we know the image set is open in $\mathbb{R}^k$ given that $a_0^{-1} \circ a_1$ is a map class C^r

8.7 Stoke's theorem

Like many of our theorems before, we start reinmann integrals defined over simple cubes/rectangles before generalizing it other sets.

Fact 236

Denote the unit k-cube in $\mathbb{R}^k$ by

$$I^k = [0,1]^k = \underbrace{[0,1] \times \ldots \times [0,1]}_{\text{k times}}$$

Lemma 237

Let $k > 1$. Let η be a $k-1$ form defined in an open set U of $\mathbb{R}^k$ containing the unit k-cube I^k. Assume that η vanishes at all points of $\operatorname{Bd} I^k$ except possibly at points of the subset $(\operatorname{Int} I^{k-1}) \times 0$. Then

$$\int_{\operatorname{Int} I^k} d\eta = (-1)^k \int_{\operatorname{Int} I^{k-1}} b^* \eta$$

where $b : I^{k-1} \to I^k$ is the inclusion map

$$b(u_1, \ldots, u_{k-1}) = (u_1, \ldots, u_{k-1}, 0)$$

Proof. Given $1 \leq j \leq k$, let I_j denote the k-1 tuple/multi-interval

$$I_j = (1, \ldots, \widehat{j}, \ldots, k)$$

Let the typical $k-1$ form in $\mathbb{R}^k$ be the form

$$dx_{I_j} = dx_1 \wedge \ldots \wedge \widehat{dx_j} \wedge \ldots \wedge$$

THen we define η to be

$$\eta = f \, dx_{I_j}$$

Then we have

$$\begin{aligned}
d\eta &= df \wedge dx_{I_j} \\
&= \left(\sum_{i=1}^{k} D_i f \, dx_i \right) \wedge dx_{I_j} \\
&= (-1)^{j-1} (D_j f) dx_1 \wedge \ldots \wedge dx_k
\end{aligned}$$

Let us first calculate the LHS

$$\begin{aligned}
\int_{\operatorname{Int} I^k} d\eta &= (-1)^{j-1} \int_{\operatorname{Int} I^k} D_j f \\
&= (-1)^{j-1} \int_{I^k} D_j f \\
&= (-1)^{j-1} \int_{v \in I^{k-1}} \int_{x_j \in I} D_j f(x_1, \ldots, x_k)
\end{aligned}$$

where the last line follows by **fubini theorem**. Then by the fundamental theorem of calculus we have

$$\int_{x_j \in I} D_j f(x_1, \ldots, x_k) = f(x_1, \ldots, 1, \ldots, x_k) - f(x_1, \ldots, 0, \ldots, x_k)$$

However by assumption, f vanishes at all points of $\operatorname{Bd} I^k$ so the RHS is zero when $j < k$, unless $j = k$ in which case the RHS is on (Int $I^{k-1} \times 0$) so it is possibly non-zero. If it is non-zero then it will take the value

$$-f(x_1, \dots, x_{k-1}, 0)$$

We therefore can conclude

$$\int_{\text{Int } I^k} d\eta = \begin{cases} 0 & j < k \\ (-1)^{k-1} \int_{I^{k-1}} -f(x_1, \dots, x_{k-1}, 0) = (-1)^k \int_{I^{k-1}} (f \circ b)(x_1, \dots, x_{k-1}) & j = k \end{cases}$$

Now let us calculate the RHS recall 163

$$b^*(dx_{I_j}) = \det \begin{bmatrix} a_{p,q} \end{bmatrix} dx_{I_k} = [\det \begin{bmatrix} \frac{\partial b_{I_j}}{\partial x_{I_k}} \end{bmatrix}] dx_1 \wedge \dots \wedge d_{x_{k-1}}$$
$$= \begin{cases} 0 & j < k \\ dx_1 \wedge \dots \wedge dx_{k-1} & j = k \end{cases}$$

where p, q are pth and qth elements of I_j and I_k respectively because

$$b((x_{I_j})_p) = \begin{cases} x_p & p \neq j \\ 0 & p = j \end{cases}$$

Hence so

$$\frac{\partial b((x_{I_j})_p)}{\partial x_q} = \begin{cases} 0 & p \neq q \\ 0 & p = q \quad \text{and} \quad p = j \\ 1 & p = q \quad \text{and} \quad p \neq j \end{cases}$$

So the $(k-1) \times (k-1)$ matrix is only the indentity when it is not possible for $p = j$, that is when $j = k$ where $x_{I_j} = x_{I_k} = (x_1, \dots, x_{k-1}, 0)$

$$\det \begin{bmatrix} \frac{\partial b_{I_j}}{\partial x_{I_k}} \end{bmatrix} = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases}$$

Finally we have

$$\int_{\text{Int } I^{k-1}} b^* \eta = \int_{\text{Int } I^{k-1}} b^* f dx_{I_j} = \begin{cases} 0 & j < k \\ \int_{I^{k-1}} (f \circ b)(x_1, \dots, x_{k-1}) & j = k \end{cases}$$

Now multiply both sides our result here and notice we get the LHS as calculated from before. Therefore the theorem follows

$$\int_{\text{Int } I^k} d\eta = (-1)^k \int_{\text{Int } I^{k-1}} b^* \eta = (-1)^k \int_{I^{k-1}} (f \circ b)(x_1, \dots, x_{k-1})$$

□

Theorem 238 (Stoke theorem)
Let $k > 1$. Let M be a compact oriented k-manifold in $\mathbb{R}^n$ give ∂M the induced orientiation if ∂M is nnot empty. Let w be a k-1 form defined in an open set of $\mathbb{R}^n$ containing M. Then

$$\int_M dw = \int_{\partial M} w$$

if ∂M is not empty and $\int_M dw = 0$ if ∂M is empty

Consider the arbitrary coordinate path $a : W \to Y$. Let $W = (\text{Int } I^k) \cup (\text{Int } I^{k-1} \times 0)$. Basically W is a set open in $\mathbb{H}^k$ recall 233. Now define a special $k-1$ form w such that the set

$$C = M \cap \text{supp } w$$

can be covered by every coordinate patch about points $p \in M$ defined by a. Now consider

$$\int_M dw = \int_{\partial M} dw + \int_{M - \partial M} dw$$

and

$$\int_{\partial M} dw = \int_{\partial M \cap Y} dw$$

$$\int_{M-\partial M} dw = \int_{(M-\partial M)\cap Y} dw$$

$$\int_M dw = \int_{M \cap Y} dw$$

because by definition w and hence dw vanishes outside Y so all these follow recall 58. Now consider

$$\int_{M-\partial M} dw = \int_{\text{Int } I^k} a^* dw \tag{1}$$

$$= \int_{\text{Int } I^k} d\eta \tag{2}$$

$$= (-1)^k \int_{\text{Int } I^{k-1}} b^* \eta \tag{3}$$

(1) follows from 235, sets open in $\mathbb{H}^k$ not of the form $\mathbb{R}^{k-1} \times 0$ will be mapped exclusively to interior points on M. And $(M - \partial M) \cap Y = \text{Int } I^k$ Next, because M $w \in \Omega_c^{n-1}(M)$ and M is connected this integral defined also makes sense by pointcare lemma. Making use 237 since we clearly already have the necessary conditions, let $a^* w = \eta$ and as usual b be the inclusion map. Then (2) and (3) naturally follows.Now define the coordinate patch/parameterization

$$\beta = a \circ b : \text{Int } I^{k-1} \to I^{k-1} \times 0 \to Y \cap \partial M$$

The idea is that we know that $b^*\eta = b^*(a^* w) = \beta^* w$. Hence we can continue from (3) to define

$$= (-1)^k \int_{\text{Int } I^{k-1}} \beta^* w$$

$$= \int_{\partial M} w$$

It can be seen that should ∂M be empty then $b^*\eta$ vanishes so the integral in (5) is zero and we get

$$\int_M dw = \int_{\partial M} dw + \int_{M-\partial M} dw = 0$$

Remark 239. *Notice in (4) that we have implicity assumed k is even. If k is odd we have to reverse in the sign in which case we have*

$$\int_M dw = -\int_{\partial M} w$$

when ∂M non empty and

$$\int_M dw = 0$$

which is still unchanged for the case when ∂M empty

8.8 orientiable manifold

Definition 240

Let V be an n-dimensional vector space. An n-tupe $(a_1, \dots, a_n)$ of indepdent vectors in V is called an **n-frame** in V. In $\mathbb{R}^n$ we call such a frame **right-handed** if

$$\det \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix} > 0$$

and **left-handed** otherwise. The collection of all right-handed frames in $\mathbb{R}^n$ is called an orientation of $\mathbb{R}^n$ More generally choose a linear isomorphism $T : \mathbb{R}^n \to V$ and define one orientation of V to consist of all frames of the form

$$(T(a_1), \dots, T(a_n))$$

for which $(a_1, \dots, a_n)$ is right-handed in $\mathbb{R}^n$

Theorem 241

Let C be a non-singular $n \times n$ matrix. Let $h : \mathbb{R}^n \to \mathbb{R}^n$ be the linear transofmration $h(x) = C \cdot x$. Let $(a_1, \dots, a_n)$be a frame in $\mathbb{R}^n$. If $\det C > 0$ the frames

$$(a_1, \dots, a_n) \quad \text{and} \quad (h(a_1), \dots, h(a_n))$$

belongn to the same orientation in $\mathbb{R}^n$. If $\det C < 0$ they belonng to opposite orientations of $\mathbb{R}^n$. The we say h is **orientation preserving** and **orientation reversing** in the former and latter case respectively.

Proof.

$$C \cdot [a_1 \ \dots \ a_n] = [b_1 \ \dots \ b_n]$$

now take det on all sides.

Definition 242

Let $g : A \to B$ be a diffeomorphism of open sets in $\mathbb{R}^k$ we say that g is **orientation preserving** if $\det Dg > 0$ and **orientation reversing** if $\det Dg < 0$

Definition 243

We sa the coordinate patches $a_i : U_i \to V_i$ on M **overlap positively** if this transition fuction $a_1^{-1} \circ a_0$ is orientation preserving. If M can be covered by a collection of coordinate patches and each pair overlaps positively then M is said to be **orientable** and otherwise **non-orientable**

Definition 244

Let M be a k-manifold in $\mathbb{R}^n$. Suppose M is orientable. Given a collection of coordinate patches covering M that overlap positively adjoin to this collection all other coordinate patches on M that overlap these patches positively. Then this expanded collection is called an **orientation** of M. A manifold M togehter with an orientation of M is called an **oriented manifold**.

It is easy to this expanded collection of coordinate patches also overlap with each other positively on M.

Definition 245

Firt recall that the **tangent space** T_pM is the space of tangent vectors at specific point p on a manifold. Let M be an oriented 1-manifold in $\mathbb{R}^n$. The **unit tangent vector field** T on M maps a each p to one tangent vector of unit lenght in T_pM. More precisely it is defined as follows: given $p \in M$ chose a coordinate patch $a : U \to V$ on M about p belonging to the given orientation of M

$$T(p) = (p; Da(t_0)/\,||Da(t_0)||) \subset T_p(M)$$

where t_0 is the parameter value such that $a(t_0) = p$. Note that $(p; Da(t_0))$ is the **velocity vector** of the curve a corresponding to the parameter value $t = t_0$.

It is well defined just consider

$$Da(t_0) = D(\beta \circ g)(t_0) = D\beta(t_1) \cdot Dg(t_0)$$

and since g is orientation preserving $Dg(t_0) > 0$

$$Da(t_0)\,||Da(t_0)|| = D\beta(t_1)/\,||D\beta(t_1)||$$

so it is independent of choice of paramterization

Definition 246

Let M be an $n-1$ manifold in $\mathbb{R}^n$. If $p \in M$ let $(p; n)$ be a unit vector i the n-dimensional vector space $T_p(\mathbb{R}^n)$ that is orthogonal to the $n-1$ dimensional linear subspace $T_p(M)$. Then n is uniquely determined p to sign. Given a coordinate patch $a : U \to V$ on M about p belonging to this orientation, let $a(x) = p$. Then the following columns of matrix $Da(x)$ give the basis

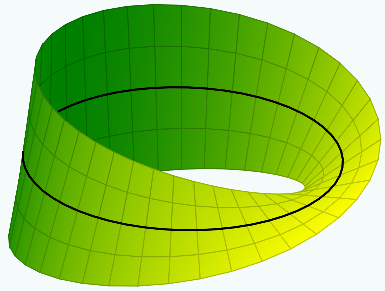
$$(p; \frac{\partial a}{\partial x_1}), \dots, (p; \frac{\partial a}{\partial x_{n-1}})$$

we specify the sign of n requiring that the frame $(n, \frac{\partial a}{\partial x_1}, \dots, \frac{\partial a}{\partial x_{n-1}})$ be right handed, that is the corresponding matrix $[n\ Da(x)]$ has positive determinant. The vector field $N(p) = (p; n(p))$ is called the **unit normal field** to M corresponding to the orientation of M.

You could see $\{\frac{\partial a}{\partial x_1}, \dots, \frac{\partial a}{\partial x_{n-1}}\}$ as the set of all tangent curves which span $T_p(M)$ and that **n** is orthogonal to all of them.

Example 247

the **mobius strip** is an example of *non-orientable* manifold



To understand more refer to `18.950 Differential Geometry(do Carmo)`

Definition 248

Let M be an n-manifod in $\mathbb{R}^n$. If $a : U \to V$ is coordinate patch on M then Da is an $n \times n$ matrix. We define the **natural orientation** of M to consist of all cooridnate patches on M for which $\det Da > 0$.

It is easy to see that these patches overlap positively. $\det {a_1}^{-1} \circ a_0 = \det {a_1}^{-1} \det a_0 > 0$

Theorem 249

Let $k > 1$. If M is an orientable k-manifold with non-empty boundary then ∂M is orientable.

Proof. a and β are coordinate pathces about p $U \to V$. Their restirctions $a|\partial\mathbb{H}^k$ an $\beta|\partial\mathbb{H}^k$ are defined by

$$a = a \circ b \quad \text{and} \quad \beta_0 = \beta \circ b$$

where $b(x_1, \ldots, x_{k-1}) = (x_1, \ldots, x_{k-1}, 0)$ is the inclusion map. We show that if a, β overlap positively then so do their resitrctions to a_0 and β_0. Let $g : W_0 \to W_1$ be the transition function $g = \beta^{-1} \circ a$ where W_0 and W_1 are open in $\mathbb{H}^k$. THen $\det Dg > 0$. Now if $x \in \partial\mathbb{H}^k$ then the derivative Dg of g at x has the last row

$$Dg_k = [0 \ \ldots \ 0 \ \frac{\partial g_K}{\partial x_k}]$$

where $\frac{\partial g_k}{\partial x_K} \geq 0$. That is because g being a diffeomrphism must satisfy

$$g : \mathbb{H}^k \to \mathbb{H}^k \quad \Leftrightarrow \quad \begin{cases} g_k(x_1, \ldots, x_n) < 0 & x_k < 0 \\ g_k(x_1, \ldots, x_k) > 0 & x_k > 0 \\ g_k(x_1, \ldots, x_k) = 0 \end{cases}$$

the first 2 follows becuase half planes must be mapped to half planes have the kth coordinate ≥ 0 by definition which are defined by the kth coordinate being greater than equal to zero. The last case follows because boundary points are mapped to boundary points. Therefore

$$\begin{cases} \frac{\partial g_k}{\partial x_k}(x_1, \ldots, x_k) \geq 0 \\ \frac{\partial g_k}{\partial x_k}(x_1, \ldots, x_k) = 0 & i \neq k \end{cases}$$

Because $\det Dg > 0$ it follows that the restriction to the boundary

$$\det \frac{\partial g_1, \ldots, g_{k-1}}{\partial x_1, \ldots, x_{k-1}} > 0$$

Therefore given an orientation of M one can obtain an orientation of ∂M by simply taking of coordinate patches that belong the to orientation of M. However this orientation of ∂M is not always the one we prefer thus we make the following definitions

Definition 250

Let M be an orientable k-manifold with non-empty boundary. Given an orientation of M the corresponding **induced orientation** of ∂M is defined as follows. If k is even it is the orientation obtained by simply taking restrictions of coordinate patches belonging to the orientation of M like above. If k is odd it is the *opposite* of the orientation of ∂M obtained in this way.

See munkres page 290 more info. The big idea is our choice of orientation by definition is where (N, W, T) is right handed where N is the unit normal vector and T is the unit tangent vector and W is a vector tangent to both.(You can see how it specifies because if you swap any the sign changes according to the number of permutations) Then use your right hand with thum pointing in N fingers to T and palm to W. Then your palm should always be pointing towards the manifold. Unfortunately i don't think munkres will go any deeper than this. Probably have to go into the realms of `algebraic topology`. Sounds like hodge star operator honestly

Fact 251 (Reversing the orientation of a manifold)

Let $r : \mathbb{R}^k \to \mathbb{R}^k$ be the reflection map

$$r(x_1, \ldots, x_k) = (-x_1, x_2, \ldots, x_k)$$

it is its own inverse as you can see. The map r carries $\mathbb{H}^k$ to $\mathbb{H}^k$ if $k > 1$ and it carries $\mathbb{H}^1$ to the left half line $\mathbb{L}^1$ if $k = 1$

Definition 252

Let M be an oriented k-manifold in $\mathbb{R}^n$. If $a_i : U_i \to V_i$ is a coordinate patch on M belonging to the orientation of M let β_i be the coordinate patch

$$\beta_i = a_i \circ r : r(U_i) \to V_i$$

Then β_i overlaps a_i negatively so it does not belong to the orientation of M. However the coordinate patches β_i overlap each other positively so they constitute a separate orientation of M. Just consider

$$Da \circ r = \frac{\partial(-x_1, x_2 \ldots x_n)}{\partial x_l} \quad \text{and} \quad Da = \frac{\partial(x_1, \ldots x_n)}{\partial x_l}$$

you can see their determinants will be of opposite polarity as the first row of each is scaled by -1 relative its equivalent. You can also tell That their respective determinants of transition maps $\det \beta_0^{-1}\beta_1 > 0$ and $\det a_0^{-1} a_1 > 0$ are positive only with respect to themselves while maps like

$$\det \beta_0^{-1} a_1 < 0 \quad \text{and} \quad \det \alpha_0^{-1} \beta_1 < 0$$

β_i is called the **reverse/opposite** orientation specified by a_i.

If M is connected it only has two. Else it has more than two.

Fact 253

recall that tensor spaces, tangent spaces, vector spaces, take the form $(p; v)$ where p is the "root" and the in linear in v. That is $(p, av_1 + bv_2) = a(p; v_1) + b(p; v_2)$

Example 254

Let M be an oriented 1 manifold with corresponding tangent field T. Then reversing the orientation of M results in replacing T with $-T$. For if $a : U \to V$ is a coordinate patch belonging to the orientation of M, then $a \circ r$ belongs to the opposite direction since

$$(a \circ r)(t) = a(-t)$$

so that

$$d(a \circ r)/dt = -da/dt = -Da$$

So

$$T(p) = (p; Da(t_0)/\,||Da(t_0)||) \quad \Rightarrow \quad -T(p) = (p; -Da(t_0)/\,||Da(t_0)||)$$

Example 255

Now let M be an $n-1$ oriented manifold with corresponding normal field N. Reversing the orientation of M results in replacing N with $-N$ since if $a : U \to V$ belongs the orientation of M , then $a \circ r$ belongs to the opposite direction since

$$\frac{\partial a \circ r}{\partial x_1} = -\frac{\partial a}{\partial x_1} \quad \text{and} \quad \frac{\partial a \circ r}{\partial x_i} = \frac{\partial a}{\partial x_i}$$

Therefore one of the frames

$$(\mathbf{n}, \frac{\partial a}{\partial x_1}, \ldots, \frac{\partial a}{\partial x_{n-1}}) \quad \text{and} \quad (-\mathbf{n}, -\frac{\partial a}{\partial x_1}, \ldots, \frac{\partial a}{\partial x_{n-1}})$$

is right-handed(as required by the definition of unit normal tangent vector) if and only if the other one is since they in fact have the same determinant value. So $\mathbf{n}$ must correspond to the coordinate pathc $a \circ r$. Recall `Artin Algebra 1` for elementary column matrix operations if you scale one column by -1 so the determinant will also be scaled by -1. So

$$N(p) = (p; \mathbf{n}) \quad \Rightarrow \quad -N(p) = (p; -\mathbf{n})$$

Definition 256

Let M be a k manifold in $\mathbb{R}^n$ let $p \in M$. If M is oriented then the taent space to M at p has a natural induced orientation defined as follows. Choose a coordinate patch $a : U \to V$ belonging to the orientation of M about p. Let $a(x) = p$. The collection of k-frames in $T_p(M)$ of the form

$$(a_*(x; a_1), \ldots, a_*(x; a_k))$$

where $(a_1, \ldots, a_k)$ is a right hnaded frame in $\mathbb{R}^k$ is called the **natural orientation** of $T_p(M)$ induced by the orientation of M.

9 vector & scalar fields

Fact 257

We denote the **gradient** by

$$\nabla \quad \text{or} \quad \text{grad}$$

The **divergence** by

$$\nabla \cdot \quad \text{or} \quad \text{div}$$

and the **curl** by

$$\nabla \times \quad \text{or} \quad \text{curl}$$

Definition 258

Let A be open in $\mathbb{R}^n$. Let $f : A \to \mathbb{R}$ be a scalar field in A. We define a corresponding vector field in A called the **gradient** of f by the equation

$$(\text{grad } f) = (x; D_1 f(x) e_1 + \ldots + D_n f(x) e_n$$

If $G(x) = (x; g(x))$ is a vector field in A where $g : A \to \mathbb{R}^n$ is given by

$$g(x) = g_1(x) e_1 + \ldots + g_n(x) e_n$$

Then we define a corresponding scalar field in A called the **divergence** of G by the equation

$$(\text{div } G)(x) = D_1 g_1(x) + \ldots + D_n g_n(x)$$

Theorem 259 (Scalar and Vector Field commutative diagram in $A \in \mathbb{R}^n$)

As usual $\Omega^k(A)$ refer to k-forms on A and d is the differential operator Let A be an open set in $\mathbb{R}^n$. There exists vector space isomorphisms a_i and β_i as in the following diagrams

$$\begin{tikzcd} \text{Scalar fields in A} \arrow[r, "a_0"] \arrow[d, "\text{grad}"'] & \Omega^0(A) \arrow[d, "d"] \\ \text{Vector fields in A} \arrow[r, "a_1"] & \Omega^1(A) \end{tikzcd}$$

and

$$\begin{tikzcd} \text{Vector fields in A} \arrow[r, "\beta_{n-1}"] \arrow[d, "\text{div}"'] & \Omega^{n-1}(A) \arrow[d, "d"] \\ \text{Scalar fields in A} \arrow[r, "\beta_n"] & \Omega^n(A) \end{tikzcd}$$

that is to say

$$d \circ a_0 = a_1 \circ \text{grad} \quad \text{and} \quad d \circ \beta_{n-1} = \beta_n \circ \text{div}$$

Proof. Let f and h be scalar fields in A. Let

$$F(x) = (x; \sum f_i(x) e_i)$$

and

$$G(x) = (x; \sum g_i(x)e_i)$$

be vector fields in A. Define the transomrations a_i and β_j by the equations

$$a_0 f = f$$

$$a_1 F = \sum_{i=1}^{n} f_i dx_i$$

$$\beta_{n-1} G = \sum_{i=1}^{n} (-1)^{i-1} g_i dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n$$

$$\beta_n h = h dx_1 \wedge \ldots \wedge dx_n$$

To show $d \circ a_0 = a_1 \circ \text{grad}$:

$$d \circ a_0(f) = df$$

$$\text{grad}(f) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} e_i$$

$$a_1(\text{grad}(f)) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i = df$$

Thus, $d \circ a_0 = a_1 \circ \text{grad}$.

To show $d \circ \beta_{n-1} = \beta_n \circ \text{div}$:

$$\beta_{n-1}(G) = \sum_{i=1}^{n} (-1)^{i-1} g_i dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n$$

$$d(\beta_{n-1}(G)) = \sum_{i=1}^{n} (-1)^{i-1} dg_i \wedge dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n$$

$$\text{div}(G) = \sum_{i=1}^{n} \frac{\partial g_i}{\partial x_i}$$

$$\beta_n(\text{div}(G)) = \left(\sum_{i=1}^{n} \frac{\partial g_i}{\partial x_i} \right) dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n$$

Since

$$dg_i = \sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j} dx_j,$$

we have

$$d(\beta_{n-1}(G)) = \sum_{i=1}^{n} \frac{\partial g_i}{\partial x_i} dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n = \beta_n(\text{div}(G))$$

Thus, $d \circ \beta_{n-1} = \beta_n \circ \text{div}$.

Fact 260

To remember all these just remember. a_0 maps scalar field to zero form. But that is basically an identity map if you recall from earlier. a_1 maps a vector field to a 1 form. Thats easy, simply assign every component e_i of the vector a dx_i. For a_{n-1} that's a little complicated, you wanna to map $dx_1 \wedge \ldots \widehat{dx_i} \ldots \wedge dx_n$ to every component of the vector field. But you want to have an $(-1)^{i-1}$ which corresponds to the number of swaps to return dx_i back upon appending it at the front. Recall that we have used this form before to prove poincare lemma before. Finally a_n is pretty easy, scalar field now just add dx_I to the back.

Definition 261

Let A be open in $\mathbb{R}^3$ let

$$F(x) = (x; \sum f_i(x)e_i)$$

be a vector field in A. We define another vector field in A called the curl of F by the equation

$$(\text{curl } F)(x) = (x; (D_2 f_3 - D_3 f_2)e_1 + (D_3 f_1 - D_1 f_3)e_2 + (D_1 f_2 - D_2 f_1)e_3$$

Theorem 262 (Scalar and Vector Field commutative diagram in $A \in \mathbb{R}^n$)

Let A be open in $\mathbb{R}^3$.There exist vector space isomorphisms a_i and β_j as in the following diagram

$$\begin{tikzcd}
\text{Scalar fields in A} \arrow[r, "a_0"] \arrow[d, "\text{grad}"'] & \Omega^0(A) \arrow[d, "d"] \\
\text{Vector fields in A} \arrow[r, "a_1"] \arrow[d, "\text{curl}"'] & \Omega^1(A) \arrow[d, "d"] \\
\text{Vector fields in A} \arrow[r, "\beta_2"] \arrow[d, "\text{div}"'] & \Omega^2(A) \arrow[d, "d"] \\
\text{Scalar fields in A} \arrow[r, "\beta_3"] & \Omega^3(A)
\end{tikzcd}$$

The new relationships between the maps are:

$$d \circ a_0 = a_1 \circ \text{grad},$$

$$d \circ a_1 = \beta_2 \circ \text{curl},$$

$$d \circ \beta_2 = \beta_3 \circ \text{div}\,.$$

Proof. The 1st statement has statement has already been proven previously.
Next, for a vector field $F(x) = \sum_{i=1}^{3} f_i(x)e_i$ in $\mathbb{R}^3$, the curl is given by:

$$\text{curl } F(x) = (D_2 f_3 - D_3 f_2)\, e_1 + (D_3 f_1 - D_1 f_3)\, e_2 + (D_1 f_2 - D_2 f_1)\, e_3.$$

The isomorphism $a_1(F) = f_1 dx_1 + f_2 dx_2 + f_3 dx_3$ maps F to a 1-form.Recall the action of β_{n-1} is as follows

$$\beta_{3-1} G = \sum_{i=1}^{3} (-1)^{i-1} g_i dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_3$$

Hence the 2-form associated with curl F under the map β_2 is:

$$\beta_2(\operatorname{curl} F) = (D_2 f_3 - D_3 f_2)\, dx_2 \wedge dx_3 + (D_3 f_1 - D_1 f_3)\, dx_3 \wedge dx_1 + (D_1 f_2 - D_2 f_1)\, dx_1 \wedge dx_2.$$

Computing the exterior derivative of $a_1(F)$,

$$d(a_1(F)) = d\left(f_1 dx_1 + f_2 dx_2 + f_3 dx_3\right) = \sum_{i=1}^{3} \sum_{j=1}^{3} D_j f_i\, dx_j \wedge dx_i,$$

we find after permutating to get matching orders and removing repeating tensors we have 3!

$$d(a_1(F)) = (D_2 f_3 - D_3 f_2)\, dx_2 \wedge dx_3 + (D_3 f_1 - D_1 f_3)\, dx_3 \wedge dx_1 + (D_1 f_2 - D_2 f_1)\, dx_1 \wedge dx_2,$$

which matches $\beta_2(\operatorname{curl} F)$. Thus,

$$d \circ a_1(F) = \beta_2 \circ \operatorname{curl}(F).$$

Finally, for a vector field $G(x) = \sum_{i=1}^{n} g_i(x) e_i$, the divergence is given by:

$$\operatorname{div} G(x) = D_1 g_1(x) + \ldots + D_n g_n(x).$$

Recall the action of β_n

$$\beta_3 h = h dx_1 \wedge \ldots \wedge dx_3$$

so

$$\beta_3(\operatorname{div} G) = (D_1 g_1(x) + \ldots + D_n g_n(x)) dx_1 \wedge dx_2 \wedge dx_3$$

Hence the isomorphism $\beta_2(G)$ gives the 2-form as:

$$\beta_2(G) = g_1 dx_2 \wedge dx_3 + g_2 dx_3 \wedge dx_1 + g_3 dx_1 \wedge dx_2,$$

The exterior derivative of $\beta_2(G)$ is:

$$d(\beta_2(G)) = d\left(g_1 dx_2 \wedge dx_3 + g_2 dx_3 \wedge dx_1 + g_3 dx_1 \wedge dx_2\right),$$

yielding

$$d(\beta_2(G)) = \sum_{i=1}^{3} D_i g_i\, dx_1 \wedge dx_2 \wedge dx_3.$$

This matches $\beta_3(\operatorname{div} G)$, so

$$d \circ \beta_2(G) = \beta_3 \circ \operatorname{div}(G).$$

Thus, the relationships between the maps are:

$$d \circ a_0 = a_1 \circ \operatorname{grad}, \quad d \circ a_1 = \beta_2 \circ \operatorname{curl}, \quad d \circ \beta_2 = \beta_3 \circ \operatorname{div}.$$

Fact 263 (Elementary Multivar Calculus Results)
Green theorem
$$\int\int_A (D_1 f_2 - D_2 f_1)\, dxdy = \oint_{\partial A} (f_1 dx + f_2 dy)$$
Classical Stoke Theorem
$$\int_A ((D_2 f_3 - D_3 f_2)\, dx_2 dx_3 + (D_3 f_1 - D_1 f_3)\, dx_3 dx_1 + (D_1 f_2 - D_2 f_1)\, dx_1 dx_2)$$
$$= \int_{\partial A} f_1 dx_1 + f_2 dx_2 + f_3 dx_3$$
All these follow by stokes theorem on 2-manifolds in $\mathbb{R}^n$
$$\int_M dw = \int_{\partial M} w$$
where the 1-form is $w = \sum_i f_i dx_i$. Then the results follow after directly calculating dw. To see how classical Stoke's theorem is related to curl and volumes, consider the proof of divergence theorem below

9.1 vector fields

Definition 264
Let U be an open set of $\mathbb{R}^n$. A **vector field** v on U is a function which assigns to each point $p \in U$ a corresponding $v(p) \in T_p(U)$. That is to say 1 point p to 1 tangent vector in T_pU. In constrast recall a **one-form** on U is a function which assigns to each point $p \in U$ an element $w(p) \in T^*_p(\mathbb{R}^n)$

Example 265
Let U be an open set of $\mathbb{R}^n$ and $p \in U$. Let e_i be the standard basis vectors of U. Then the vector field $p \to (p; e_i)$ which we denote as $\left(\frac{\partial}{\partial x_i}\right)_p$ if you recall. These tuple pairs ar cearly the basis of T_pU. So we can write every vector field as
$$v(p) = \sum_i f_i(p) \left(\frac{\partial}{\partial x_i}\right)_p = \sum_i f_i(p)(p; e_i) \quad \Rightarrow \quad v = \sum f_i \frac{\partial}{\partial x_i}$$
where f_i are real valued function($\Omega^0(U)$) - coefficients of our basis vectors

Definition 266
The **tangent space** $T_p(M)$ is essentially
$$T_p(M) = a_*(T_x(\mathbb{R}^k))$$
where a is a coordinate patch. It is clear to see that $T_x(\mathbb{R}^k)$ is spanned by (p, e_i) while $T_p(M)$ is spanned by $(p, Da(x) \cdot e_i) = (p, \frac{\partial a}{\partial x_i})$

Definition 267

Let U be an open set in $\mathbb{R}^n$ and let v be a *vector field* on U. A function $\gamma : (a, b) \to U$ is an integral curve of v if for all $a < t < b$ if we define $p = \gamma(t)$ we have

$$v(p) = \left(p, \frac{d\gamma}{dt}(t)\right)$$

Definition 268

A vector field v on U is **complete** if for every $p \in U$ there exists an integral curve $\gamma_p(t) : \mathbb{R} \to U$ such that $\gamma_p(0) = p$ ad the map $F : U \times (-\infty, \infty) \to U$ defined by $F(p, t) = \gamma_p(t)$ is a C^∞ map.

Proposition 269

If a vector field v is *compactly supported* then v is **complete**

Proof. Suppose we have $v(p_0) = 0$. THen the curve $\gamma_0(t) = p_0$ for $t \in \mathbb{R}$ satisfies

$$0 = \frac{d}{dt}\gamma(t) = v(p_0)$$

..to be continued

9.2 tensor fields

Definition 270

A **k-tensor field** in A is a function assigning to each $x \in A$ a k-tensor defined on the vector space $T_x(\mathbb{R}^n)$ that is

$$w(x) \in \mathcal{L}^k(T_x(\mathbb{R}^n))$$

if it so happens that $w(x)$ is an alternating for each x that is

$$w(x) \in \mathcal{L}^k(T_x(\mathbb{R}^n)) = \Lambda^k(T_x^*(\mathbb{R}^n))$$

then we say w is **k-form**(recall above)

9.3 Divergence Theorem

Another reminder that ds(1-dimensional version of dV, the displacement) and dV are *distinct* from the differential operator dx_i. They way they are calculated is by $\int f\,dV = \int_{Y_a}(f \circ v)V(Da)$ where V is a special function certainly not calculated in the same way as how v is associated with integrals with respect to dx_i. Their equivalence needs to be considered on a case by case basis which we will now proceed to do so below.

Fact 271 (Informal Overview 1)
The big idea for the below theorems is that given a a *vector field* $F = \sum f_i e_i$ or $G = \sum g_i e_i$ if the corresponding form is related by

$$a_1 F = \sum_{i=1}^{n} f_i dx_i = w \quad \to \quad \int_M w = \int_M \langle F, T \rangle ds$$

where w is a 1 form defined on an 1 manifold. So projection of F onto the unit tangent field over the line integral. The "rough idea" is that the magnitude of the tangent vector defines an infinitesimal so if we integrate over the line we the get the "length" of the curve

$$\beta_{n-1} G = \sum_{i=1}^{n} (-1)^{i-1} g_i dx_1 \wedge \ldots \widehat{dx_i} \ldots \wedge dx_n = w \quad \to \quad \int_M w = \int_M \langle G, N \rangle dV$$

where w is an $n-1$ form defined on an $n-1$ manifold. So projection of G onto the unit normal field over the surface integral. A "rough idea" why this works is that the magnitude of the unit normal defines a infinitesimal small area at the surface so if we integrate over the surface we the get the "volume" of the surface. In fact you could see the latter case as the higher dimensional version of the former since a 1D area is analagous to lenght. We will prove more rigorously below.

Example 272 (Gradient Theorem)
A special case when our vector field is $F = \operatorname{grad} f$ on a 1 manifold we have

$$a_1 \operatorname{grad} f = \sum_{i=1}^{n} D_i f \, dx_i = df = w \quad \to \quad \int_M df = \int_M \langle \operatorname{grad} F, T \rangle ds$$

Fact 273 (Informal Overview 2)
On the other hand, given a *scalar field* f the corresponding form is related by

$$a_0 f = f = w \quad \to \quad \int w = \int f$$

where w is an 0 form. This is clear as zero form and scalar field are analogous.

$$\beta_n f = f \, dx_1 \wedge \ldots \wedge dx_n = w \quad \to \quad \int w = \int f dV$$

where w is an n form defined on an $n-1$ manifold. You could see how this make sense as $\int_{I^k} f \, dx_1 \ldots dx_n$ where I_n is n-multi interval that defines the domain we are integrating with respect to. Then the multiplication of all the ranges f is integrated over gives f multiplied by the volume in a sense. Again we will prove more rigorously below.

Example 274 (Divergence Theorem)
Given G but in the special case where corresponding form is $w = \beta_{n-1}G$ is instead defined on the boundary of an $n-1$ manifold. Then we have

$$\int_{\partial M} w = \int_{\partial M} \langle G, N\rangle dV$$

In which we case using stokes theorem which states that

$$\int_{\partial M} w = \int_M dw$$

we can make an additional simplification knowing that $dw = \beta_n \circ \operatorname{div} G$ since $d \circ \beta_{n-1} = \beta_n \circ \operatorname{div}$ we can do

$$\int_M dw = \int_M (\operatorname{div} G) dV$$

and so we have the divergence theorem

$$\int_M (\operatorname{div} G) dV = \int_{\partial M} \langle G, N\rangle dV$$

As a reminder just see ds as a 1-dimensional dV

Example 275 (Classical Stokes Theorem)
Given F but in the special case where corresponding form is $w = \alpha_1 F$ is instead defined on the boundary of an 1 manifold. Then we have

$$\int_{\partial M} w = \int_{\partial M} \langle F, T\rangle ds$$

In which we case using stokes theorem which states that

$$\int_{\partial M} w = \int_M dw$$

we can make an additional simplification knowing that $dw = \beta_2 \circ \operatorname{curl} F$ since $d \circ \alpha_1 = \beta_2 \circ \operatorname{curl}$ we can do

$$\int_M dw = \int_M \langle \operatorname{curl} F, N\rangle dV$$

and so we have the classical stokes theorem

$$\int_M \langle \operatorname{curl} F, N\rangle dV = \int_{\partial M} \langle F, T\rangle ds$$

We will now develop the proofs below

Lemma 276 (line integral)
Let M be a compact oriented 1-manifold in $\mathbb{R}^n$ let T be the unit tangent vector to M corresponding to the otrientation. Let

$$F(x) = (x; f(x)) = (x; \sum f_i(x)e_i)$$

be a vector field in an open set o $\mathbb{R}^n$ containing M which corresponds to the 1-form

$$w = \sum f_i dx_i$$

Then

$$\int_M w = \int_M \langle F, T\rangle ds$$

where ds is the 1-dimensional notation of dV

Proof. Let the set

$$C = M \cap \operatorname{supp} w$$

lie in a single coordinate patch $a : U \to V$ belonging to orientatin of M. In that case we have

$$\begin{aligned} a^* w &= \sum_{i=1}^{n} (f_i \circ a) d(x_i \circ a) = \sum_{i=1}^{n} (f_i \circ a) da_i \\ &= \sum_{i=1}^{n} (f_i \circ a)(Da_i) dt \\ &= \langle f \circ a, Da\rangle dt \end{aligned}$$

where the second line follows by change of variables $\frac{\partial a_i}{\partial t} = Da_i$. So it follows that

$$\int_M w = \int_U a^* w = \int_U \langle f \circ a, Da\rangle$$

On the other hand

$$\int_M \langle F, T\rangle ds = \int_U \langle F \circ a, T \circ a\rangle \cdot V(Da) = \int_U \langle f \circ a, Da/\,||Da||\rangle \cdot V(Da) = \int_U \langle f \circ a, Da\rangle$$

since recall 180 and 232

$$V(Da) = [\det((Da)^t Da)]^{\frac{1}{2}} = ||Da||$$

so by the linearity of inner products knowing that $||Da||$ is a 1D real number/scalar the lemma follows

Definition 277
Let M be an oriented $n-1$ manifold in $\mathbb{R}^n$. Given $p \in M$ let $(p; n)$ be a unit vector in $T_p(\mathbb{R}^n)$ that is orthogonal to the $n-1$ dimensional linear subspace $T_p(M)$. If $a : U \to V$ is a coordinate patch on M about p belonging to the orientation of M(orientation preserving parameterization) with $a(x) = p$ choose n so that

$$\left(n, \frac{\partial a}{\partial x_1}(x), \ldots, \frac{\partial a}{\partial x_{n-1}}(x)\right)$$

is right handed meaning the determinant of that frame above is greater than zero. Then the vector fied $N(p) = (p; n(p))$ is called the **unit normal field** corresponding to the orientation of M.

Proposition 278

Show that $N(p)$ is well defined and of class C^∞

Proof. To show it is well defined we need to show it is indepedent of paramterization function like how we have done previously. As usual we let $g = \beta^{-1} \circ \alpha$ be the transtion function and we see that by chain rule

$$Da(x) = D\beta(y) \cdot Dg(x)$$

where $g(x) = y$ then calculating the determinant we have

$$\det[v \ D\alpha(x)] = \det[v \ D\beta(y)] \cdot \det Dg(x)$$

Since $Dg > 0$(g is a composition of orientation preserving functions) we conclude $[v \ D\alpha(x)]$ has positive determinant if and only if $[v \ D\beta(y)]$ does. To show N is C^∞ consider its definition

$$N(p) = (p; c(x)/\,||c(x)||)$$

where $x = a^{-1}(p)$ and $c = \sum c_i(x)e_i$. So it is essentially a linear map which is smooth as shown in 169. The motivation for c is formulated in the following lemma.

Lemma 279

Given independent vectors $x_1, \ldots, x_{n-1}$ in $\mathbb{R}^n$ let X be the $n \times n-1$ matrix $X = [x_1, \ldots, x_{n-1}]$ and let c be the vector $c = \sum c_i e_i$ where

$$c_i = (-1)^{i-1} \det \underbrace{X(1, \ldots, \widehat{i}, \ldots, n)}_{\text{exclude ith row}}$$

The vector c has the following properties

1. c is non-zero and orthogonal to each x_i
2. The frame $(c, x_1, \ldots, x_{n-1})$ is right-handed
3. $||c|| = V(X)$

Remark 280. *Quite a smart way to define tangent fields. We essentially use the determinant function to define it by letting c_i be the cofactors of the matrix. So whatever column vector c is gonna inherit properties of the determinant. The last property could be interpreted as the "volume" of the unit tangent space.*

Proof. Fix $x_1, \ldots, x_{n-1}$. Given $a \in \mathbb{R}^n$ compute the following determinant expanding by the cofactors of the 1st column we have

$$\det[a \ x_1 \ \ldots \ x_{n-1}] = \sum_{i=1}^{n} a_i(-1)^{i+1} \det X(1, \ldots, \widehat{i}, \ldots, n) = \sum_{i=1}^{n} a_i(-1)^{i-1} \det X(1, \ldots, \widehat{i}, \ldots, n) = \langle a, c \rangle$$

For (1) it is easy to see $\langle a, c \rangle = 0$ if a equal any x_i since will have repeated columns. Next it can be seen that both the row and column space of

$$X(1, \ldots, \widehat{i}, \ldots, n)$$

are spanned by $n-1$ vectors. Hence $\det X(1, \ldots, \widehat{i}, \ldots, n) \neq 0$ so $c_i \neq 0$ since $e_i \neq 0$ as they are the standard basis vectors. So $c \neq 0$. For (2) consider

$$\det[c\ x_1 \ldots\ x_{n-1}] = \langle c, c \rangle = ||c||^2 > 0$$

For (3) since orthogonal we have

$$[c\ x_1 \ldots\ x_{n-1}]^t [c\ x_1 \ldots\ x_{n-1}] = \begin{bmatrix} ||c||^2 & 0 \\ 0 & X^t X \end{bmatrix}$$

Taking determinants here and knowing from 2 that $\det[c\ x_1 \ldots\ x_{n-1}] = ||c||^2$ we have

$$\det \begin{bmatrix} ||c||^2 & 0 \\ 0 & X^t X \end{bmatrix} = ||c||^2 V(X)^2 = \det^2[c\ x_1 \ldots\ x_{n-1}] = ||c||^4$$

since $||c|| \neq 0$ we have $||c|| = V(X)$

Lemma 281 (surface integral)

Let M be a compact oriented $n-1$ manifold in $\mathbb{R}^n$ and let N be the corresponding unit normal vector field. Let G be a vector field defined in an open set U of $\mathbb{R}^n$ containing M. If we denote the general point of $\mathbb{R}^n$ by y and that this vector field has the form

$$G(y) = (y; g(y)) = (y; \sum g_i(y) e_i)$$

which corresponds to the $n-1$ form

$$w = \sum_{i=1}^{n} (-1)^{i-1} g_i dy_1 \wedge \ldots \wedge \widehat{dy_i} \wedge \ldots \wedge dy_n$$

Then

$$\int_M w = \int_M \langle G, N \rangle dV$$

Proof. Again it suffices to prove the theorem in the case where the set

$$C = M \cap \operatorname{supp} w$$

lies in a single coordinate path $a : U \to V$ belonging to the orientation of M. Then

$$\int_M w = \int_U a^* w = \int_U \sum_{i=1}^{n} (-1)^{i-1} (g_i \circ a) Da(1, \ldots, \widehat{i}, \ldots, n)$$

by 163 since $Da = \left[\frac{\partial a_i}{\partial x_j}\right]$. Then letting

$$c_i = (-1)^{i-1} \det \underbrace{X(1, \ldots, \widehat{i}, \ldots, n)}_{\text{exclude ith row}}$$

where $c = \sum c_i e_i$ then like we have done previously we obtain the unit normal field

$$N \circ a = N(a(x)) = (a(x); c(x)/\,||c(x)||)$$

Then we compute

$$\begin{aligned}\int_M \langle G, N\rangle dV &= \int_U \langle G\circ a, N\circ a\rangle \cdot V(Da)\\ &= \int_U \langle G\circ a, c\rangle\\ &= \int_U \sum_{i=1}^{n}(g_i\circ a)(-1)^{i-1}\det Da(1,\dots,\widehat{i},\dots,n)\end{aligned}$$

where the second equality is due to the fact that $||c|| = V(Da)$ so we may simplify like so by the linearity of inner products. So the lemma follows

Lemma 282 (volume form)
Let M be a compact n-manifold in $\mathbb{R}^n$ oriented naturally. Let $w = hdx_1\wedge\dots\wedge dx_n$ be an n-form defined in an open set of $\mathbb{R}^n$ containing M. Then h is the corresponding scalar field and

$$\int_M w = \int_M hdV$$

Proof. Again consider a similar coordinate patch $a : U \to V$ and by definition we have

$$\int_M w = \int_U a^* w = \int_U (h\circ a)\det Da$$

and

$$\int_M hdV = \int_U (h\circ a)V(Da)$$

but because orientation preserving $V(Da) = |\det Da| = \det Da$ s the lemma follows

Lemma 283
Let M be an n-manifold in $\mathbb{R}^n$. If M is oriented naturally then the induced orientation of ∂M corresponds to the uni normal field N to ∂M that points outwards from M at each point of ∂M. The **inward normal** to ∂M at p is the velocity vector of a curve that begins at p and moves into M as the parameter value increases. The **outward normal** is its negative.

Proof. Again consider a similar coordiniate patch as above on M about p belonging to orientation of M. That is $\det Da > 0$. Let $b : \mathbb{R}^{n-1}\to\mathbb{R}^n$ be te map

$$b(x_1,\dots,x_{n-1}) = (x_1,\dots,x_{n-1},0)$$

The map $a_0 = a\circ b$ is coordinate pathc on ∂M about p. It belongs to the induced orientation of ∂M if n is even and to its oppoiste if n is odd. Let N be the unit normal field to ∂M corresponding to the induced orientation of ∂M. Let $N(p) = (p; \mathbf{n}(p))$. Then the below follows by 255

$$\det\Big[(-1)^n\mathbf{n}\ Da_0\Big] > 0$$

but this implies

$$\det\Big[Da_0\ \mathbf{n}\Big] = \det\Big[\frac{\partial a}{\partial x_1}\dots\frac{\partial a}{\partial x_{n-1}}\ \mathbf{n}\Big] < 0$$

Again this because we carried out $n-1$ column swaps to get there so this follows by the properties of determinants.

However because we have assumed a natural orientation of M we have

$$\det Da = \det \left[\frac{\partial a}{\partial x_1} \dots \frac{\partial a}{\partial x_{n-1}} \; \frac{\partial a}{\partial x_n} \right] > 0$$

By comparison we see that the polarity of $\frac{\partial a}{\partial x_n}$ and $\mathbf{n}$ are opposite. So if we know that $\frac{\partial a}{\partial x_n}$ is the velocity vector of the curve that egins at a point of ∂M and moves into M as the paramter increaseses then $\mathbf{n}$ must be outward normal to ∂M at p.

Theorem 284 (Divergence Theorem)

Let M be a compact n-manifold in $\mathbb{R}^n$. Let N be the unit normal vector field to ∂M that points outwards from M. If G is a vector field defined in an open set of $\mathbb{R}^n$ then

$$\int_M (\operatorname{div} G) dV = \int_{\partial M} \langle G, N \rangle dV$$

Proof. Given G, let $w = \beta_{n-1} G$ be the corresponding $n-1$ form. Orient M naturally and give ∂M the induced orientation. Then the normal field N corresponds to the orientation of ∂M by the previous lemma. By 281 we have

$$\int_{\partial M} w = \int_{\partial M} \langle G, N \rangle dV$$

where

$$w = \sum_{i=1}^{n} (-1)^{i-1} g_i dy_1 \wedge \dots \wedge \widehat{dy_i} \wedge \dots \wedge dy_n$$

However according to 259 $dw = \beta_n(\operatorname{div} G) = (\operatorname{div} G) dx_1, \wedge \dots \wedge dx_n$. Therefore by 282 we have

$$\int_M dw = \int_M (\operatorname{div} G) dx_1 \wedge \dots \wedge dx_n = \int_M (\operatorname{div} G) dV$$

But by **stokes theorem** we have

$$\int_{\partial M} dw = \int_M dw$$

so therefore the divergence theorem follows. ☐

We now apply our results to special cases in $\mathbb{R}^3$ where we can now use the curl relationship as you can recall.

Theorem 285 (Stokes Theorem for 2-manifolds in $\mathbb{R}^3$)

Let M be a compact 2-manifold in $\mathbb{R}^3$. Let N be a unit normal field to M. Let F be a C^∞ vector field defined in an open set about M. If ∂M is empty then

$$\int_M \langle \operatorname{curl} F, N \rangle dV = 0$$

and if ∂M non empty then let T be the unit tangent vector field to ∂M chosen so that the vector $W(p) = N(p) \times T(p)$ points into M from ∂M(recall this our default choice of orientation). Then

$$\int_M \langle \operatorname{curl} F, N \rangle dV = \int_{\partial M} \langle F, T \rangle ds$$

Proof. Given the vector field $F = \sum f_i e_i$. Let $a_1 F$ be the corresponding 1-form.

$$w = a_1 F = f_1 dx_1 + f_2 dx_2 + f_3 dx_3$$

Then according to 262 the vector field curl F corresponds to the 2 form dw since

$$\beta_2 \operatorname{curl} F = d(a_1 F) = dw$$

then by 282

$$\int_M dw = \int_M \langle \operatorname{curl} F, N \rangle dV$$

evalutating the RHS of our proposition using 276 we have

$$\int_{\partial M} w = \int_{\partial M} \langle F, T \rangle ds$$

Then the theorem now follow from **stokes theorem**

Remark 286. *This was the classical version of stokes theorem before the generalized version*

9.4 interior product

Definition 287

Let V be an n-dimensional vector space and let $T \in \mathcal{L}^k(V)$. The **interior product operation** is defined as follows. For $v \in V, \iota_v(T)$ is the (k-1)tensor given by

$$\iota_v(T)(v_1, \ldots, v_{k-1}) = \sum_{r=1}^{k} (-1)^{r-1} T(v_1, \ldots, v_{r-1}, v, v_{r+1}, \ldots, v_{k-1})$$

for alternating vectors where w is *k-form* we have

$$\iota(v)w = \sum_{r=1}^{n} (-1)^{r-1} \ell_r(v) \ell_1 \wedge \ldots \wedge \ell_{r-1} \wedge \ell_{r+1} \wedge \ldots \wedge \ell_k$$

Proposition 288

Suppose $w \in \Lambda^k(V^*)$. The interior product operation which sends it $\Lambda^{k-1}(V^*)$ satisfies

1. $\iota_{v_1+v_2} w = \iota_{v_1} w + \iota_{v_2}$
2. $\iota_v(w_1 \wedge w_2) = \iota_v w_1 \wedge w_2 + (-1)^p w_1 \wedge \iota_v w_2$
3. $\iota_v(\iota_v w) = 0$

Proof. For (1): Let's compute $\iota_{v_1+v_2}(T)$ using the definition:

$$\iota_{v_1+v_2} T(v_1, \ldots, v_{k-1}) = \sum_{r=1}^{k} (-1)^{r-1} T(v_1, \ldots, v_{r-1}, v_1 + v_2, v_{r+1}, \ldots, v_{k-1})$$

By the linearity of T in each argument, this expands to:

$$\iota_{v_1+v_2} T(v_1, \ldots, v_{k-1}) = \sum_{r=1}^{k} (-1)^{r-1} \left(T(v_1, \ldots, v_{r-1}, v_1, v_{r+1}, \ldots, v_{k-1}) + T(v_1, \ldots, v_{r-1}, v_2, v_{r+1}, \ldots, v_{k-1}) \right)$$

This can be split into two sums:

$$\iota_{v_1+v_2}T(v_1,\dots,v_{k-1}) = \sum_{r=1}^{k}(-1)^{r-1}T(v_1,\dots,v_{r-1},v_1,v_{r+1},\dots,v_{k-1}) + \sum_{r=1}^{k}(-1)^{r-1}T(v_1,\dots,v_{r-1},v_2,v_{r+1},\dots,v_{k-1})$$

By definition, this is:

$$\iota_{v_1+v_2}T = \iota_{v_1}T + \iota_{v_2}T$$

For (2): Let w_1 and w_2 be decomposable p-form and q-form respectively:

$$w_1 = \ell_1 \wedge \ell_2 \wedge \dots \wedge \ell_p, \quad w_2 = \ell_{p+1} \wedge \ell_{p+2} \wedge \dots \wedge \ell_{p+q}$$

Then,

$$w_1 \wedge w_2 = \ell_1 \wedge \ell_2 \wedge \dots \wedge \ell_p \wedge \ell_{p+1} \wedge \ell_{p+2} \wedge \dots \wedge \ell_{p+q}$$

Applying the interior product:

$$\iota_v(w_1 \wedge w_2) = \sum_{r=1}^{p+q}(-1)^{r-1}\ell_r(v)\,\ell_1 \wedge \dots \wedge \ell_{r-1} \wedge \ell_{r+1} \wedge \dots \wedge \ell_{p+q}$$

This sum can be split into two parts: the sum where r ranges over the indices $1,\dots,p$ and the sum where r ranges over $p+1,\dots,p+q$:

$$\iota_v(w_1\wedge w_2) = \sum_{r=1}^{p}(-1)^{r-1}\ell_r(v)\,\ell_1\wedge\dots\wedge\ell_{r-1}\wedge\ell_{r+1}\wedge\dots\wedge\ell_p\wedge w_2 + (-1)^p w_1\wedge\sum_{s=1}^{q}(-1)^{s-1}\ell_{p+s}(v)\,\ell_{p+1}\wedge\dots\wedge\ell_{p+s-1}\wedge\ell_{p+s+1}\wedge\dots\wedge\ell_{p+q}$$

The first sum is precisely $\iota_v w_1 \wedge w_2$, and the second sum is $(-1)^p w_1 \wedge \iota_v w_2$. Thus,

$$\iota_v(w_1 \wedge w_2) = \iota_v w_1 \wedge w_2 + (-1)^p w_1 \wedge \iota_v w_2$$

For (3): Consider the interior product $\iota_v(\iota_v w)$ for some k-form w:

$$\iota_v(\iota_v w)(v_1,\dots,v_{k-2}) = \sum_{r=1}^{k}(-1)^{r-1}\iota_v w(v_1,\dots,v_{r-1},v,v_{r+1},\dots,v_{k-2})$$

Expanding $\iota_v w$ inside this expression, we have:

$$= \sum_{r=1}^{k}(-1)^{r-1}\sum_{s=1}^{k-1}(-1)^{s-1}w(v_1,\dots,v_{s-1},v,v_{s+1},\dots,v_{r-1},v,v_{r+1},\dots,v_{k-2})$$

This expression is zero because each term involves a repeated vector v (since w is antisymmetric, having a repeated argument means the whole term vanishes). Thus,

$$\iota_v(\iota_v w) = 0$$

9.5 Lie Differentiation

Definition 289

Let $w \in \Omega^k(U)$ and let v be a C^∞ vector field. The **lie differentiation** operation is defined via

$$\mathcal{L}_v w = \iota(v)dw + d(\iota(v)w)$$

Proposition 290

THe lie differentiation operations commutes with the d operation

$$\mathcal{L}_v dw = d\mathcal{L}_v w$$

Additionally when interacting with the wedge product we have

$$\mathcal{L}_v(w_1 \wedge w_2) = \mathcal{L}_v w_1 \wedge w_2 + w_1 \wedge \mathcal{L}_v w_2$$

References

[1] Prof. Victor Guillemin. *18.101 | Fall 2005 | Undergraduate Analysis II OCW Lecture Notes*. 2005. URL: https://ocw.mit.edu/courses/18-101-analysis-ii-fall-2005/resources/lectures/.