

Contents

1	Review of Topology	3
1.1	Metric Spaces	3
1.2	Open and Closed Sets	3
1.3	Metrics on \mathbb{R}^n	4
1.4	Continuity	5
1.5	Limit Points and Closure	6
1.6	Compactness	7
1.7	Connectedness	9
2	Differentiation	11
2.1	Differentiation in n dimensions	11
2.2	Conditions for Differentiability	15
2.3	Chain Rule	19
2.4	The Mean-value Theorem in n Dimensions	21
2.5	Inverse Function Theorem	22
3	Integration	29
3.1	Riemann Integral of One Variable	29
3.2	Riemann Integral of Several Variables	31
3.3	Conditions for Integrability	33
3.4	Fubini Theorem	40
3.5	Properties of Riemann Integrals	43
3.6	Integration Over More General Regions	47
3.7	Improper Integrals	49
3.8	Exhaustions	53
3.9	Support and Compact Support	57
3.10	Partitions of Unity	58
4	Multi-linear Algebra	63
4.1	Review of Linear Algebra and Topology	63
4.2	Dual Space	64
4.3	Tensors	66
4.4	Pullback Operators	69
4.5	Alternating Tensors	69
4.6	Wedge Product	80
4.7	Determinant	85
4.8	Orientations of Vector Spaces	87
4.9	Tangent Spaces and k -forms	89
4.10	Pullback Operation on Exterior Forms	96

5	Integration with Differential Forms	100
5.1	The Poincare Lemma	104
5.2	Proper Maps and Degree	111
5.3	Topological Invariance of Degree	123
6	Manifolds	127
6.1	Canonical Submersion and Canonical Immersion Theorems	127
6.2	Definition of Manifold	129
6.3	Examples of Manifolds	130
6.4	Tangent Spaces of Manifolds	134
6.5	Differential Forms on Manifolds	138
6.6	Orientation of Manifolds	142
6.7	Integration on Manifolds	146
6.8	Degree on Manifolds	150
6.9	Hopf Theorem	154
6.10	Integration on Smooth Domains	156
6.11	Stokes' Theorem	161

Lecture 1

1 Review of Topology

1.1 Metric Spaces

Definition 1.1. Let X be a set. Define the Cartesian product $X \times X = \{(x, y) : x, y \in X\}$.

Definition 1.2. Let $d : X \times X \rightarrow \mathbb{R}$ be a mapping. The mapping d is a *metric on* X if the following four conditions hold for all $x, y, z \in X$:

- (i) $d(x, y) = d(y, x)$,
- (ii) $d(x, y) \geq 0$,
- (iii) $d(x, y) = 0 \iff x = y$, and
- (iv) $d(x, z) \leq d(x, y) + d(y, z)$.

Given a metric d on X , the pair (X, d) is called a *metric space*.

Suppose d is a metric on X and that $Y \subseteq X$. Then there is an automatic metric d_Y on Y defined by restricting d to the subspace $Y \times Y$,

$$d_Y = d|_{Y \times Y}. \tag{1.1}$$

Together with Y , the metric d_Y defines the automatic metric space (Y, d_Y) .

1.2 Open and Closed Sets

In this section we review some basic definitions and propositions in topology. We review open sets, closed sets, norms, continuity, and closure. Throughout this section, we let (X, d) be a metric space unless otherwise specified.

One of the basic notions of topology is that of the *open set*. To define an open set, we first define the ϵ -neighborhood.

Definition 1.3. Given a point $x_o \in X$, and a real number $\epsilon > 0$, we define

$$U(x_o, \epsilon) = \{x \in X : d(x, x_o) < \epsilon\}. \tag{1.2}$$

We call $U(x_o, \epsilon)$ the ϵ -neighborhood of x_o in X .

Given a subset $Y \subseteq X$, the ϵ -neighborhood of x_o in Y is just $U(x_o, \epsilon) \cap Y$.

Definition 1.4. A subset U of X is *open* if for every $x_o \in U$ there exists a real number $\epsilon > 0$ such that $U(x_o, \epsilon) \subseteq U$.

We make some propositions about the union and intersections of open sets. We omit the proofs, which are fairly straightforward.

The following Proposition states that arbitrary unions of open sets are open.

Proposition 1.5. *Let $\{U_\alpha, \alpha \in I\}$ be a collection of open sets in X , where I is just a labeling set that can be finite or infinite. Then, the set*

$$\bigcup_{\alpha \in I} U_\alpha \text{ is open.}$$

The following Corollary is an application of the above Proposition.

Corollary 1. *If $Y \subset X$ and A is open in Y (w.r.t. d_Y), then there exists an open set U in X such that $U \cap Y = A$.*

Proof. The set A is open in Y . So, for any $p \in A$ there exists an $\epsilon_p > 0$ such that $U(p, \epsilon_p) \cap Y \subseteq A$. We construct a set U containing A by taking the union of the sets $U(p, \epsilon_p)$ over all p in A ,

$$U = \bigcup_{p \in A} U(p, \epsilon_p). \quad (1.3)$$

For every $p \in A$, we have $U(p, \epsilon_p) \cap Y \subseteq A$, which shows that $U \cap Y \subseteq A$. Furthermore, the union is over all $p \in A$, so $A \subseteq U$, which implies that $A \subseteq U \cap Y$. This shows that $U \cap Y = A$. To conclude the proof, we see that U is open by the openness of the $U(p, \epsilon_p)$ and the above theorem. \square

The following Proposition states that finite intersections of open sets are open.

Proposition 1.6. *Let $\{U_i, i = 1, \dots, N\}$ be a finite collection of open sets in X . Then the set*

$$\bigcap_{i=1}^{i=N} U_i \text{ is open.}$$

Definition 1.7. Define the *complement of A in X* to be $A^c = X - A = \{x \in X : x \notin A\}$.

We use the complement to define *closed* sets.

Definition 1.8. The set A is *closed in X* if A^c is open in X .

1.3 Metrics on \mathbb{R}^n

For most of this course, we will only consider the case $X = \mathbb{R}^n$ or X equals certain subsets of \mathbb{R}^n called manifolds, which we will define later.

There are two interesting metrics on \mathbb{R}^n . They are the *Euclidean metric* and the *sup metric*, and are defined in terms of the *Euclidean norm* and the *sup norm*, respectively.

Definition 1.9. Let $x \in \mathbb{R}^n$, written out in component form as $x = (x_1, x_2, \dots, x_n)$. The *Euclidean norm* of x is

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

and the *sup norm* of x is

$$|x| = \max_i |x_i|.$$

From these norms we obtain the Euclidean distance function

$$\|x - y\| \tag{1.4}$$

and the sup distance function

$$|x - y|, \tag{1.5}$$

respectively.

These two distance functions are related in several ways. In particular,

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

These distance functions are also related by the following Proposition, which will sometimes come in handy.

Proposition 1.10. *A subset U of \mathbb{R}^n is open w.r.t. the $\| \cdot \|$ distance function if and only if it is open w.r.t. the $| \cdot |$ distance function.*

So, these two distance functions give the same topologies of \mathbb{R}^n .

1.4 Continuity

Consider two metric spaces (X, d_X) and (Y, d_Y) , a function $f : X \rightarrow Y$, and a point $x_o \in X$.

Definition 1.11. The function f is *continuous at x_o* if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$d_X(x, x_o) < \delta \implies d_Y(f(x), f(x_o)) < \epsilon. \tag{1.6}$$

By definition, a function is continuous if it is continuous at all points in its domain.

Definition 1.12. The function f is *continuous* if f is continuous at every point $x_o \in X$.

There is an alternative formulation of continuity that we present here as a theorem.

Theorem 1.13. *The function f is continuous if and only if for every open subset U of Y , the pre-image $f^{-1}(U)$ is open in X .*

Continuous functions can often be combined to construct other continuous functions. For example, if $f, g : X \rightarrow \mathbb{R}$ are continuous functions, then $f + g$ and fg are continuous functions.

1.5 Limit Points and Closure

As usual, let (X, d) be a metric space.

Definition 1.14. Suppose that $A \subseteq X$. The point $x_o \in X$ is a *limit point* of A if for every ϵ -neighborhood $U(x_o, \epsilon)$ of x_o , the set $U(x_o, \epsilon)$ is an infinite set.

Definition 1.15. The *closure* of A , denoted by \bar{A} , is the union of A and the set of limit points of A ,

$$\bar{A} = A \cup \{x_o \in X : x_o \text{ is a limit point of } A\}. \quad (1.7)$$

Now we define the interior, exterior, and the boundary of a set in terms of open sets. In the following, we denote the complement of A by $A^c = X - A$.

Definition 1.16. The set

$$\text{Int } A \equiv (\bar{A}^c)^c \quad (1.8)$$

is called the *interior* of A .

It follows that

$$x \in \text{Int } A \iff \exists \epsilon > 0 \text{ such that } U(x, \epsilon) \subset A. \quad (1.9)$$

Note that the interior of A is open.

We define the exterior of a set in terms of the interior of the set.

Definition 1.17. The *exterior* of A is defined to be $\text{Ext } A \equiv \text{Int } A^c$.

The boundary of a set is the collection of all points not in the interior or exterior.

Definition 1.18. The *boundary* of A is defined to be $\text{Bd } A \equiv X - ((\text{Ext } A) \cup (\text{Int } A))$.

Always, we have $X = \text{Int } A \cup \text{Ext } A \cup \text{Bd } A$.

Lecture 2

1.6 Compactness

As usual, throughout this section we let (X, d) be a metric space. We also remind you from last lecture we defined the open set

$$U(x_o, \lambda) = \{x \in X : d(x, x_o) < \lambda\}. \quad (1.10)$$

Remark. If $U(x_o, \lambda) \subseteq U(x_1, \lambda_1)$, then $\lambda_1 > d(x_o, x_1)$.

Remark. If $A_i \subseteq U(x_o, \lambda_i)$ for $i = 1, 2$, then $A_1 \cup A_2 \subseteq U(x_o, \lambda_1 + \lambda_2)$.

Before we define compactness, we first define the notions of boundedness and covering.

Definition 1.19. A subset A of X is *bounded* if $A \subseteq U(x_o, \lambda)$ for some λ .

Definition 1.20. Let $A \subseteq X$. A collection of subsets $\{U_\alpha \subseteq X, \alpha \in I\}$ is a *cover* of A if

$$A \subset \bigcup_{\alpha \in I} U_\alpha.$$

Now we turn to the notion of compactness. First, we only consider compact sets as subsets of \mathbb{R}^n .

For any subset $A \subseteq \mathbb{R}^n$,

$$A \text{ is compact} \iff A \text{ is closed and bounded.}$$

The above statement holds true for \mathbb{R}^n but not for general metric spaces. To motivate the definition of compactness for the general case, we give the Heine-Borel Theorem.

Heine-Borel (H-B) Theorem. *Let $A \subseteq \mathbb{R}^n$ be compact and let $\{U_\alpha, \alpha \in I\}$ be a cover of A by open sets. Then a finite number of U_α 's already cover A .*

The property that a finite number of the U_α 's cover A is called the Heine-Borel (H-B) property. So, the H-B Theorem can be restated as follows: If A is compact in \mathbb{R}^n , then A has the H-B property.

Sketch of Proof. First, we check the H-B Theorem for some simple compact subsets of \mathbb{R}^n . Consider rectangles $Q = I_1 \times \cdots \times I_n \subset \mathbb{R}^n$, where $I_k = [a_k, b_k]$ for each k . Starting with one dimension, it can be shown by induction that these rectangles have the H-B property.

To prove the H-B theorem for general compact subsets, consider any closed and bounded (and therefore compact) subset A of \mathbb{R}^n . Since A is bounded, there exists a rectangle Q such that $A \subseteq Q$. Suppose that the collection of subsets $\{U_\alpha, \alpha \in I\}$ is

an open cover of A . Then, define $U_o = \mathbb{R}^n - A$ and include U_o in the open cover. The rectangle Q has the H-B property and is covered by this new cover, so there exists a finite subcover covering Q . Furthermore, the rectangle Q contains A , so the finite subcover also covers A , proving the H-B Theorem for general compact subsets. \square

The following theorem further motivates the general definition for compactness.

Theorem 1.21. *If $A \subseteq \mathbb{R}^n$ has the H-B property, then A is compact.*

Sketch of Proof. We need to show that the H-B property implies A is bounded (which we leave as an exercise) and closed (which we prove here).

To show that A is closed, it is sufficient to show that A^c is open. Take any $x_o \in A^c$, and define

$$C_N = \{x \in \mathbb{R}^n : d(x, x_o) \leq 1/N\}, \quad (1.11)$$

and

$$U_N = C_N^c. \quad (1.12)$$

Then,

$$\bigcap C_N = \{x_o\} \quad (1.13)$$

and

$$\bigcup U_N = \mathbb{R}^n - \{x_o\}. \quad (1.14)$$

The U_N 's cover A , so the H-B Theorem implies that there is a finite subcover $\{U_{N_1}, \dots, U_{N_k}\}$ of A . We can take $N_1 < N_2 < \dots < N_k$, so that $A \subseteq U_{N_k}$. By taking the complement, it follows that $C_{N_k} \subseteq A^c$. But $U(x_o, 1/N_k) \subseteq C_{N_k}$, so x_o is contained in an open subset of A^c . The above holds for any $x_o \in A^c$, so A^c is open. \square

Let us consider the above theorem for arbitrary metric space (X, d) and $A \subseteq X$.

Theorem 1.22. *If $A \subseteq X$ has the H-B property, then A is closed and bounded.*

Sketch of Proof. The proof is basically the same as for the previous theorem. \square

Unfortunately, the converse is not always true. Finally, we come to our general definition of compactness.

Definition 1.23. A subset $A \subseteq X$ is *compact* if it has the H-B property.

Compact sets have many useful properties, some of which we list here in the theorems that follow.

Theorem 1.24. *Let (X, d_X) and (Y, d_Y) be metric spaces, and let $f : X \rightarrow Y$ be a continuous map. If A is a compact subset of X , then $f(A)$ is a compact subset of Y .*

Proof. Let $\{U_\alpha, \alpha \in I\}$ be an open covering of $f(A)$. Each pre-image $f^{-1}(U_\alpha)$ is open in X , so $\{f^{-1}(U_\alpha) : \alpha \in I\}$ is an open covering of A . The H-B Theorem says that there is a finite subcover $\{f^{-1}(U_{\alpha_i}) : 1 \leq i \leq N\}$. It follows that the collection $\{U_{\alpha_i} : 1 \leq i \leq N\}$ covers $f(A)$, so $f(A)$ is compact. \square

A special case of the above theorem proves the following theorem.

Theorem 1.25. *Let A be a compact subset of X and $f : X \rightarrow \mathbb{R}$ be a continuous map. Then f has a maximum point on A .*

Proof. By the above theorem, $f(A)$ is compact, which implies that $f(A)$ is closed and bounded. Let $a = \text{l.u.b. of } f(A)$. The point a is in $f(A)$ because $f(A)$ is closed, so there exists an $x_o \in A$ such that $f(x_o) = a$. \square

Another useful property of compact sets involves the notion of uniform continuity.

Definition 1.26. Let $f : X \rightarrow \mathbb{R}$ be a continuous function, and let A be a subset of X . The map f is *uniformly continuous on A* if for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$d(x, y) < \delta \implies |f(x) - f(y)| < \epsilon,$$

for all $x, y \in A$.

Theorem 1.27. *If $f : X \rightarrow Y$ is continuous and A is a compact subset of X , then f is uniformly continuous on A .*

Proof. Let $p \in A$. There exists a $\delta_p > 0$ such that $|f(x) - f(p)| < \epsilon/2$ for all $x \in U(p, \delta_p)$. Now, consider the collection of sets $\{U(p, \delta_p/2) : p \in A\}$, which is an open cover of A . The H-B Theorem says that there is a finite subcover $\{U(p_i, \delta_{p_i}/2) : 1 \leq i \leq N\}$. Choose $\delta \leq \min \delta_{p_i}/2$. The following claim finishes the proof.

Claim. *If $d(x, y) < \delta$, then $|f(x) - f(y)| < \epsilon$.*

Proof. Given x , choose p_i such that $x \in U(p_i, \delta_{p_i}/2)$. So, $d(p_i, x) < \delta_{p_i}/2$ and $d(x, y) < \delta < \delta_{p_i}/2$. By the triangle inequality we conclude that $d(p_i, y) < \delta_{p_i}$. This shows that $x, y \in U(p_i, \delta_{p_i})$, which implies that $|f(x) - f(p_i)| < \epsilon/2$ and $|f(y) - f(p_i)| < \epsilon/2$. Finally, by the triangle inequality, $|f(x) - f(y)| < \epsilon$, which proves our claim. \square

1.7 Connectedness

As usual, let (X, d) be a metric space.

Definition 1.28. The metric space (X, d) is *connected* if it is impossible to write X as a disjoint union $X = U_1 \cup U_2$ of non-empty open sets U_1 and U_2 .

Note that disjoint simply means that $U_1 \cap U_2 = \phi$, where ϕ is the empty set.

A few simple examples of connected spaces are \mathbb{R} , \mathbb{R}^n , and $I = [a, b]$. The following theorem shows that a connected space gets mapped to a connected subspace by a continuous function.

Theorem 1.29. *Given metric spaces (X, d_X) and (Y, d_Y) , and a continuous map $f : X \rightarrow Y$, it follows that*

$$X \text{ is connected} \implies f(X) \text{ is connected.}$$

Proof. Suppose $f(X)$ can be written as a union of open sets $f(X) = U_1 \cup U_2$ such that $U_1 \cap U_2 = \phi$. Then $X = f^{-1}(U_1) \cup f^{-1}(U_2)$ is a disjoint union of open sets. This contradicts that X is connected. \square

The intermediate-value theorem follows as a special case of the above theorem.

Intermediate-value Theorem. *Let (X, d) be connected and $f : X \rightarrow \mathbb{R}$ be a continuous map. If $a, b \in f(X)$ and $a < r < b$, then $r \in f(X)$.*

Proof. Suppose $r \notin f(X)$. Let $A = (-\infty, r)$ and $B = (r, \infty)$. Then $X = f^{-1}(A) \cup f^{-1}(B)$ is a disjoint union of open sets, a contradiction. \square

Lecture 3

2 Differentiation

2.1 Differentiation in n dimensions

We are setting out to generalize to n dimensions the notion of differentiation in one-dimensional calculus. We begin with a review of elementary one-dimensional calculus.

Let $I \subseteq \mathbb{R}$ be an open interval, let $f : I \rightarrow \mathbb{R}$ be a map, and let $a \in I$.

Definition 2.1. The *derivative of f at a* is

$$f'(a) = \lim_{t \rightarrow 0} \frac{f(a+t) - f(a)}{t}, \quad (2.1)$$

provided that the limit exists. If the limit exists, then f is *differentiable at a* .

There are half a dozen or so possible reasonable generalizations of the notion of derivative to higher dimensions. One candidate generalization which you have probably already encountered is the directional derivative.

Definition 2.2. Given an open set U in \mathbb{R}^n , a map $f : U \rightarrow \mathbb{R}^m$, 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

$$D_u f(a) = \lim_{t \rightarrow 0} \frac{f(a+tu) - f(a)}{t}, \quad (2.2)$$

provided that the limit exists.

In particular, we can calculate the directional derivatives in the direction of the standard basis vectors e_1, \dots, e_n of \mathbb{R}^n , where

$$e_1 = (1, 0, \dots, 0), \quad (2.3)$$

$$e_2 = (0, 1, 0, \dots, 0), \quad (2.4)$$

$$\vdots \quad (2.5)$$

$$e_n = (0, \dots, 0, 1). \quad (2.6)$$

Notation. The directional derivative in the direction of a standard basis vector e_i of \mathbb{R}^n is denoted by

$$D_i f(a) = D_{e_i} f(a) = \frac{\partial}{\partial x_i} f(a). \quad (2.7)$$

We now try to answer the following question: What is an adequate definition of differentiability at a point a for a function $f : U \rightarrow \mathbb{R}^m$?

- Guess 1: Require that $\frac{\partial f}{\partial x_i}(a)$ exists.

However, this requirement is inadequate. Consider the function defined by

$$f(x_1, x_2) = \begin{cases} 0, & \text{if } (x_1, x_2) \text{ lies on the } x_1\text{-axis or the } x_2\text{-axis,} \\ 1, & \text{otherwise.} \end{cases} \quad (2.8)$$

Then, both

$$\frac{\partial f}{\partial x_1}(0) = 0 \text{ and } \frac{\partial f}{\partial x_2}(0) = 0, \quad (2.9)$$

but the function f is not differentiable at $(0, 0)$ along any other direction.

- Guess 2: Require that all directional derivatives exist at a .

Unfortunately, this requirement is still inadequate. For example (from Munkres chapter 5), consider the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ defined by

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2+y^4}, & (x, y) \neq (0, 0) \\ 0, & x = y = 0. \end{cases} \quad (2.10)$$

Claim. *The directional derivative $D_u f(0)$ exists for all u .*

Proof. Let $u = (h, k)$. Then

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{f(tu) - f(0)}{t} &= \lim_{t \rightarrow 0} \frac{f(tu)}{t} \\ &= \lim_{t \rightarrow 0} \left(\frac{t^3 h k^2}{t^2 h^2 + t^4 k^4} \right) \frac{1}{t} \\ &= \begin{cases} 0, & h = 0 \\ k^2/h, & h \neq 0. \end{cases} \end{aligned} \quad (2.11)$$

So the limit exists for every u . □

However, the function is a non-zero constant on a parabola passing through the origin: $f(t^2, t) = \frac{t^4}{2t^4} = \frac{1}{2}$, except at the origin where $f(0, 0) = 0$. The function f is discontinuous at the origin despite the existence of all directional derivatives.

- Guess 3. This guess will turn out to be correct.

Remember that in one-dimensional calculus we defined

$$f'(a) = \lim_{t \rightarrow 0} \frac{f(a+t) - f(a)}{t}, \quad (2.12)$$

for a function $f : I \rightarrow \mathbb{R}$ and a point $a \in I$. Now consider the function $\lambda : \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$\lambda(t) = f'(a)t. \quad (2.13)$$

Then,

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{f(a+t) - f(a) - \lambda(t)}{t} &= \lim_{t \rightarrow 0} \frac{f(a+t) - f(a)}{t} - f'(a) \\ &= 0. \end{aligned} \quad (2.14)$$

So, $\lambda(t) \approx f(a+t) - f(a)$ when t is small.

Now we generalize to n dimensions.

Definition 2.3. Given an open subset U of \mathbb{R}^n , a map $f : U \rightarrow \mathbb{R}^m$, and a point $a \in U$, the function f is *differentiable at a* if there exists a linear mapping $B : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that for every $h \in \mathbb{R}^n - \{0\}$,

$$\frac{f(a+h) - f(a) - Bh}{|h|} \rightarrow 0 \text{ as } h \rightarrow 0. \quad (2.15)$$

That is, $f(a+h) - f(a) \approx Bh$ when h is small.

Theorem 2.4. *If f is differentiable at a , then for every u the directional derivative of f in the direction of u at a exists.*

Proof. The function f is differentiable at a , so

$$\frac{f(a+tu) - f(a) - B(tu)}{|tu|} \rightarrow 0 \text{ as } t \rightarrow 0. \quad (2.16)$$

Furthermore,

$$\begin{aligned} \frac{f(a+tu) - f(a) - B(tu)}{|tu|} &= \frac{t}{|tu|} \frac{f(a+tu) - f(a) - B(tu)}{t} \\ &= \frac{t}{|t|} \frac{1}{|u|} \left(\frac{f(a+tu) - f(a)}{t} - Bu \right) \\ &\rightarrow 0, \end{aligned} \quad (2.17)$$

as $t \rightarrow 0$, so

$$\frac{f(a+tu) - f(a)}{t} \rightarrow Bu \text{ as } t \rightarrow 0. \quad (2.18)$$

□

Furthermore, the linear map B is unique, so the following definition is well-defined.

Definition 2.5. The *derivative of f at a* is $Df(a) = B$, where B is the linear map defined above.

Note that $Df(a) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear map.

Theorem 2.6. *If f is differentiable at a , then f is continuous at a .*

Sketch of Proof. Note that for $h \neq 0$ in \mathbb{R}^n ,

$$\frac{f(a+h) - f(a) - Bh}{|h|} \rightarrow 0 \text{ as } h \rightarrow 0 \quad (2.19)$$

implies that

$$f(a+h) - f(a) - Bh \rightarrow 0 \text{ as } h \rightarrow 0. \quad (2.20)$$

From this you can conclude that f is continuous at a . \square

Remark. Let $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear map and $a \in \mathbb{R}^n$. The point a can be written as a sum $a = \sum_{j=1}^n a_j e_j = (a_1, \dots, a_n)$. The point La can be written as the sum $La = \sum a_j L e_j$, and L can be written out in components as $L = (L_1, \dots, L_m)$, where each $L_j : \mathbb{R}^n \rightarrow \mathbb{R}$ is a linear map. Then $L e_j = (L_1 e_j, \dots, L_m e_j)$, and $L_i e_j = \ell_{i,j}$. The numbers $\ell_{i,j}$ form an $n \times n$ matrix denoted by $[\ell_{i,j}]$.

Remark. Let $U \subseteq \mathbb{R}^n$, and let $f_1 : \mathbb{R}^n \rightarrow \mathbb{R}^{m_1}$ and $f_2 : \mathbb{R}^n \rightarrow \mathbb{R}^{m_2}$ be differentiable maps. Let $m = m_1 + m_2$, so that $\mathbb{R}^{m_1} \times \mathbb{R}^{m_2} = \mathbb{R}^m$. Now, construct a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ defined in component form by $f = (f_1, f_2)$. The derivative of f at a is

$$Df(a) = (Df_1(a), Df_2(a)). \quad (2.21)$$

Remark. Let $f : U \rightarrow \mathbb{R}^m$ be a map. The action of f on input x written out in component form is $f(x) = (f_1(x), \dots, f_m(x))$. So, the map can be represented in component form as $f = (f_1, \dots, f_m)$, where each f_i as a map of the form $f_i : U \rightarrow \mathbb{R}$. The derivative of f acting on the standard basis vector e_j is

$$\begin{aligned} Df(a)e_j &= (Df_1(a)e_j, \dots, Df_m(a)e_j) \\ &= \left(\frac{\partial f_1}{\partial x_j}(a), \dots, \frac{\partial f_m}{\partial x_j}(a) \right). \end{aligned} \quad (2.22)$$

So, the derivative $(Df)(a)$ can be represented by an $m \times n$ matrix

$$(Df)(a) \cong J_f(a) = \left[\frac{\partial f_i}{\partial x_j}(a) \right] \quad (2.23)$$

called the Jacobian matrix of f at a , which you probably recognize.

Lecture 4

2.2 Conditions for Differentiability

In this lecture we will discuss conditions that guarantee differentiability. First, we begin with a review of important results from last lecture.

Let U be an open subset of \mathbb{R}^n , let $f : U \rightarrow \mathbb{R}^m$ be a map, and let $a \in U$.

We defined f to be differentiable at a if there exists a linear map $B : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that for $h \in \mathbb{R}^n - \{0\}$,

$$\frac{f(a+h) - f(a) - Bh}{|h|} \rightarrow 0 \text{ as } h \rightarrow 0. \quad (2.24)$$

If such a B exists, then it is unique and $B = Df(a)$. The matrix representing B is the Jacobian matrix $J_f(a) = \left[\frac{\partial f_i}{\partial x_j}(a) \right]$, where $f = (f_1, \dots, f_m)$.

Note that the mere existence of all of the partial derivatives in the Jacobian matrix does not guarantee differentiability.

Now we discuss conditions that guarantee differentiability.

Theorem 2.7. *Suppose that all of the partial derivatives $\frac{\partial f_i}{\partial x_j}$ 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 .*

Sketch of Proof. This theorem is very elegantly proved in Munkres, so we will simply give the general ideas behind the proof here.

First, we look at the case $n = 2, m = 1$. The main ingredient in the proof is the Mean Value Theorem from 1-D calculus, which we state here without proof.

Mean Value Theorem. *Given an interval $[a, b] \subseteq \mathbb{R}$ and a map $\phi : [a, b] \rightarrow \mathbb{R}$, if ϕ is continuous on $[a, b]$ and differentiable on (a, b) , then there exists a point $c \in (a, b)$ such that $\phi(b) - \phi(a) = \phi'(c)(b - a)$.*

Now we continue with the proof. Let f be a map $f : U \rightarrow \mathbb{R}$, where $U \subseteq \mathbb{R}^2$. So, f is a function of two variables $f = f(x_1, x_2)$. Consider a point $a = (a_1, a_2) \in U$ and any point $h \in \mathbb{R}^2 - \{0\}$ “close” to zero, where by close we mean $a + h \in U$. We want to compute $f(a + h) - f(a)$.

$$\begin{aligned} f(a+h) - f(a) &= f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) \\ &= f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) \\ &\quad + f(a_1, a_2 + h_2) - f(a_1, a_2). \end{aligned} \quad (2.25)$$

Thinking of the first two terms as functions of the first argument only, and thinking of the last two terms as functions of the second term only, and applying the Mean

Value Theorem to each pair of terms, we obtain

$$\begin{aligned} f(a+h) - f(a) &= \frac{\partial f}{\partial x_1}(c_1, a_2 + h_2)h_1 \\ &\quad + \frac{\partial f}{\partial x_2}(a_1, d_2)h_2, \end{aligned} \tag{2.26}$$

where $a_1 < c_1 < a_1 + h_1$ and $a_2 < d_2 < a_2 + h_2$. This can be rewritten as

$$f(a+h) - f(a) = \frac{\partial f}{\partial x_1}(c)h_1 + \frac{\partial f}{\partial x_2}(d)h_2, \tag{2.27}$$

where $c = (c_1, a_2 + h_2)$ and $d = (a_1, d_2)$.

We want to show that $(f(a+h) - f(a) - Df(a)h)/|h| \rightarrow 0$ as $h \rightarrow 0$, where $Df(a) = \left[\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a) \right]$. Using our previously derived expression for $f(a+h) - f(a)$, we find that

$$\begin{aligned} f(a+h) - f(a) - Df(a)h &= f(a+h) - f(a) - \frac{\partial f}{\partial x_1}(a)h_1 - \frac{\partial f}{\partial x_2}(a)h_2 \\ &= \left(\frac{\partial f}{\partial x_1}(c) - \frac{\partial f}{\partial x_1}(a) \right) h_1 + \left(\frac{\partial f}{\partial x_2}(d) - \frac{\partial f}{\partial x_2}(a) \right) h_2. \end{aligned} \tag{2.28}$$

We can use the sup norm to show that

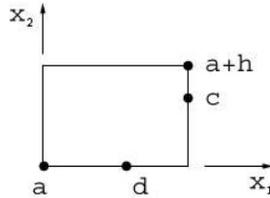
$$|f(a+h) - f(a) - Df(a)h| \leq \left| \frac{\partial f}{\partial x_1}(c) - \frac{\partial f}{\partial x_1}(a) \right| |h_1| + \left| \frac{\partial f}{\partial x_2}(d) - \frac{\partial f}{\partial x_2}(a) \right| |h_2|, \tag{2.29}$$

from which it follows that

$$\frac{|f(a+h) - f(a) - Df(a)h|}{|h|} \leq \left| \frac{\partial f}{\partial x_1}(c) - \frac{\partial f}{\partial x_1}(a) \right| + \left| \frac{\partial f}{\partial x_2}(d) - \frac{\partial f}{\partial x_2}(a) \right|, \tag{2.30}$$

where we used the fact that $|h| = \max(|h_1|, |h_2|)$.

Notice that as $h \rightarrow 0$, both $c \rightarrow a$ and $d \rightarrow a$, as can be easily seen using the following diagram. This means that the r.h.s. of Equation (2.30) goes to zero as h



goes to zero, because the partial derivatives are continuous. It follows that the l.h.s. goes to zero, which completes our proof.

The proof in n dimensions is similar to the above proof, but the details are harder to follow. \square

We now introduce a useful class of functions.

Definition 2.8. Given $U \subseteq \mathbb{R}^n$ and $f : U \rightarrow \mathbb{R}$, we define

$$f \in \mathcal{C}^1(U) \iff \frac{\partial f}{\partial x_i}, i = 1, \dots, n \text{ exist and are continuous at all points } x \in U. \quad (2.31)$$

Similarly, we define

$$f \in \mathcal{C}^2(U) \iff \frac{\partial f}{\partial x_i} \in \mathcal{C}^1(U), i = 1, \dots, n. \quad (2.32)$$

$$f \in \mathcal{C}^k(U) \iff \frac{\partial f}{\partial x_i} \in \mathcal{C}^{k-1}(U), i = 1, \dots, n. \quad (2.33)$$

$$f \in \mathcal{C}^\infty(U) \iff f \in \mathcal{C}^k(U) \forall k. \quad (2.34)$$

If f is multiply differentiable, then you can perform higher order mixed partial derivatives.

One of the fundamental theorems of calculus is that the order of the partial derivatives can be taken in any order. For example,

$$\frac{\partial}{\partial x_i} \left(\frac{\partial}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left(\frac{\partial}{\partial x_i} \right) \equiv \frac{\partial^2 f}{\partial x_i \partial x_j} \quad (2.35)$$

Let's do the proof for this case. Let $U \subseteq \mathbb{R}^2$ and $f = f(x_1, x_2)$. We prove the following claim:

Claim.

$$\frac{\partial}{\partial x_i} \left(\frac{\partial}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left(\frac{\partial}{\partial x_i} \right). \quad (2.36)$$

Proof. Take $a \in U$ written in components as $a = (a_1, a_2)$, and take $h = (h_1, h_2) \in \mathbb{R}^2 - \{0\}$ such that $a + h \in U$. That is, take $h \approx 0$.

Define

$$\Delta(h) = f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) - f(a_1 + h_1, a_2) + f(a_1, a_2), \quad (2.37)$$

and define

$$\phi(s) = f(a_1 + h_1, s) - f(a_1, s), \quad (2.38)$$

where $a_2 \leq s \leq a_2 + h_2$. We find that

$$\begin{aligned} \Delta(h) &= \phi(a_2 + h_2) - \phi(a_2) \\ &= \phi'(c_2)h_2, a_2 < c_2 < a_2 + h_2, \end{aligned} \quad (2.39)$$

by the Mean Value Theorem. Writing out ϕ' using partial derivatives of f , and using the Mean Value Theorem again, we find

$$\begin{aligned}
\Delta(h) &= \left(\frac{\partial f}{\partial x_2}(a_1 + h_1, c_2) - \frac{\partial f}{\partial x_1}(a_1, c_2) \right) h_2 \\
&= \left(\frac{\partial}{\partial x_1} \left(\frac{\partial f}{\partial x_2}(c_1, c_2) h_1 \right) \right) h_2, a_1 < c_1 < a_1 + h_1 \\
&= \left(\frac{\partial}{\partial x_1} \left(\frac{\partial}{\partial x_2} f \right) \right) (c) h_1 h_2 \\
&= \left(\frac{\partial}{\partial x_2} \left(\frac{\partial}{\partial x_1} f \right) \right) (d) h_1 h_2,
\end{aligned} \tag{2.40}$$

where we obtained the last line by symmetry. This shows that

$$\frac{\partial}{\partial x_1} \left(\frac{\partial f}{\partial x_2} \right) (c) = \frac{\partial}{\partial x_2} \left(\frac{\partial f}{\partial x_1} \right) (d). \tag{2.41}$$

As $h \rightarrow 0, c \rightarrow a$ and $d \rightarrow a$, so

$$\frac{\partial}{\partial x_1} \left(\frac{\partial f}{\partial x_2} \right) (a) = \frac{\partial}{\partial x_2} \left(\frac{\partial f}{\partial x_1} \right) (a), \tag{2.42}$$

for any $a \in U$. □

The above argument can be iterated for $f \in \mathcal{C}^3(U)$, $f \in \mathcal{C}^3(4)$, etc.

Lecture 5

2.3 Chain Rule

Let U and v be open sets in \mathbb{R}^n . Consider maps $f : U \rightarrow V$ and $g : V \rightarrow \mathbb{R}^k$. Choose $a \in U$, and let $b = f(a)$. The composition $g \circ f : U \rightarrow \mathbb{R}^k$ is defined by $(g \circ f)(x) = g(f(x))$.

Theorem 2.9. *If f is differentiable at a and g is differentiable at b , then $g \circ f$ is differentiable at a , and the derivative is*

$$(Dg \circ f)(a) = (Dg)(b) \circ Df(a). \quad (2.43)$$

Proof. This proof follows the proof in Munkres by breaking the proof into steps.

- Step 1: Let $h \in \mathbb{R}^n - \{0\}$ and $h \doteq 0$, by which we mean that h is very close to zero. Consider $\Delta(h) = f(a + h) - f(a)$, which is continuous, and define

$$F(h) = \frac{f(a + h) - f(a) - Df(a)h}{|h|}. \quad (2.44)$$

Then f is differentiable at a if and only if $F(h) \rightarrow 0$ as $h \rightarrow 0$.

$$F(h) = \frac{\Delta(h) - Df(a)h}{|h|}, \quad (2.45)$$

so

$$\Delta(h) = Df(a)h + |h|F(h). \quad (2.46)$$

Lemma 2.10.

$$\frac{\Delta(h)}{|h|} \text{ is bounded.} \quad (2.47)$$

Proof. Define

$$|Df(a)| = \sup_i \left| \frac{\partial f}{\partial x_i}(a) \right|, \quad (2.48)$$

and note that

$$\frac{\partial f}{\partial x_i}(a) = Df(a)e_i, \quad (2.49)$$

where the e_i are the standard basis vectors of \mathbb{R}^n . If $h = (h_1, \dots, h_n)$, then $h = \sum h_i e_i$. So, we can write

$$Df(a)h = \sum h_i Df(a)e_i = \sum h_i \frac{\partial f}{\partial x_i}(a). \quad (2.50)$$

It follows that

$$\begin{aligned} |Df(a)h| &\leq \sum_{i=1}^m h_i \left| \frac{\partial f}{\partial x_i}(a) \right| \\ &\leq m|h||Df(a)|. \end{aligned} \quad (2.51)$$

By Equation 2.46,

$$|\Delta(h)| \leq m|h||Df(a)| + |h|F(h), \quad (2.52)$$

so

$$\frac{|\Delta(h)|}{|h|} \leq m|Df(a)| + F(h). \quad (2.53)$$

□

- Step 2: Remember that $b = f(a)$, $g : V \rightarrow \mathbb{R}^k$, and $b \in V$. Let $k \doteq 0$. This means that $k \in \mathbb{R}^n - \{0\}$ and that k is close to zero. Define

$$G(k) = \frac{g(b+k) - g(b) - (Dg)(b)k}{|k|}, \quad (2.54)$$

so that

$$g(b+k) - g(b) = Dg(b)k + |k|G(k). \quad (2.55)$$

We proceed to show that $g \circ f$ is differentiable at a .

$$\begin{aligned} g \circ f(a+h) - g \circ f(a) &= g(f(a+h)) - g(f(a)) \\ &= g(b + \Delta(h)) - g(b), \end{aligned} \quad (2.56)$$

where $f(a) = b$ and $f(a+h) = f(a) + \Delta(h) = b + \Delta(h)$. Using Equation 2.55 we see that the above expression equals

$$Dg(b)\Delta(h) + |\Delta(h)|G(\Delta(h)). \quad (2.57)$$

Substituting in from Equation 2.46, we obtain

$$\begin{aligned} g \circ f(a+h) - g \circ f(a) &= \dots \\ &= Dg(b)(Df(a)h + |h|F(h)) + \dots \\ &= Dg(b) \circ Df(a)h + |h|Dg(b)F(h) + |\Delta(h)|G(\Delta(h)) \end{aligned} \quad (2.58)$$

This shows that

$$\frac{g \circ f(a+h) - g \circ f(a) - Dg(b) \circ Df(a)h}{|h|} = Dg(b)F(h) + \frac{\Delta(h)}{|h|}G(\Delta(h)). \quad (2.59)$$

We see in the above equation that $g \circ f$ is differentiable at a if and only if the l.h.s. goes to zero as $h \rightarrow 0$. It suffices to show that the r.h.s. goes to zero as $h \rightarrow 0$, which it does: $F(h) \rightarrow 0$ as $h \rightarrow 0$ because f is differentiable at a ; $G(\Delta(h)) \rightarrow 0$ because g is differentiable at b ; and $\Delta(h)/|h|$ is bounded.

□

We consider the same maps g and f as above, and we write out f in component form as $f = (f_1, \dots, f_n)$ where each $f_i : U \rightarrow \mathbb{R}$. We say that f is a \mathcal{C}^r map if each $f_i \in \mathcal{C}^r(U)$. We associate $Df(x)$ with the matrix

$$Df(x) \sim \left[\frac{\partial f_i}{\partial x_j}(x) \right]. \quad (2.60)$$

By definition, f is \mathcal{C}^r (that is to say $f \in \mathcal{C}^r(U)$) if and only if Df is \mathcal{C}^{r-1} .

Theorem 2.11. *If $f : U \rightarrow V \subseteq \mathbb{R}^n$ is a \mathcal{C}^r map and $g : V \rightarrow \mathbb{R}^p$ is a \mathcal{C}^r map, then $g \circ f : U \rightarrow \mathbb{R}^p$ is a \mathcal{C}^r map.*

Proof. We only prove the case $r = 1$ and leave the general case, which is inductive, to the student.

- Case $r = 1$:

$$Dg \circ f(x) = Dg(f(x)) \circ Df(x) \sim \left[\frac{\partial g_i}{\partial x_j} f(x) \right]. \quad (2.61)$$

The map g is \mathcal{C}^1 , which implies that $\partial g_i / \partial x_j$ is continuous. Also,

$$Df(x) \sim \left[\frac{\partial f_i}{\partial x_j} \right] \quad (2.62)$$

is continuous. It follows that $Dg \circ f(x)$ is continuous. Hence, $g \circ f$ is \mathcal{C}^1 .

□

2.4 The Mean-value Theorem in n Dimensions

Theorem 2.12. *Let U be an open subset of \mathbb{R}^n and $f : U \rightarrow \mathbb{R}$ a \mathcal{C}^1 map. For $a \in U$, $h \in \mathbb{R}^n$, and $h \neq 0$,*

$$f(a+h) - f(a) = Df(c)h, \quad (2.63)$$

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

Proof. Define a map $\phi : [0, 1] \rightarrow \mathbb{R}$ by $\phi(t) = f(a+th)$. The Mean Value Theorem implies that $\phi(1) - \phi(0) = \phi'(c) = (Df)(c)h$, where $0 < c < 1$. In the last step we used the chain rule. □

2.5 Inverse Function Theorem

Let U and V be open sets in \mathbb{R}^n , and let $f : U \rightarrow V$ be a \mathcal{C}^1 map. Suppose there exists a map $g : V \rightarrow U$ that is the inverse map of f (which is also \mathcal{C}^1). That is, $g(f(x)) = x$, or equivalently $g \circ f$ equals the identity map.

Using the chain rule, if $a \in U$ and $b = f(a)$, then

$$(Dg)(b) = \text{the inverse of } Df(a). \quad (2.64)$$

That is, $Dg(b) \circ Df(a)$ equals the identity map. So,

$$Dg(b) = (Df(a))^{-1} \quad (2.65)$$

However, this is not a trivial matter, since we do not know if the inverse exists. That is what the inverse function theorem is for: if $Df(a)$ is invertible, then g exists for some neighborhood of a in U and some neighborhood of $f(a)$ in V . We state this more precisely in the following lecture.

Lecture 6

We begin with a review of some earlier definitions.

Let $\delta > 0$ and $a \in \mathbb{R}^n$.

$$\text{Euclidean ball: } B_\delta(a) = \{x \in \mathbb{R}^n : \|x - a\| < \delta\} \quad (2.66)$$

$$\begin{aligned} \text{Supremum ball: } R_\delta(a) &= \{x \in \mathbb{R}^n : |x - a| < \delta\} \\ &= I_1 \times \cdots \times I_n, \quad I_j = (a_j - \delta, a_j + \delta). \end{aligned} \quad (2.67)$$

Note that the supremum ball is actually a rectangle. Clearly, $B_\delta(a) \subseteq R_\delta(a)$. We use the notation $B_\delta = B_\delta(0)$ and $R_\delta = R_\delta(0)$.

Continuing with our review, given U open in \mathbb{R}^n , a map $f : U \rightarrow \mathbb{R}^k$, and a point $a \in U$, we defined the derivate $Df(a) : \mathbb{R}^n \rightarrow \mathbb{R}^k$ which we associated with the matrix

$$Df(a) \sim \left[\frac{\partial f_i}{\partial x_j}(a) \right], \quad (2.68)$$

and we define

$$|Df(a)| = \sup_{i,j} \left| \frac{\partial f_i}{\partial x_j}(a) \right|. \quad (2.69)$$

Lastly, we define $U \subseteq \mathbb{R}^n$ to be *convex* if

$$a, b \in U \implies (1-t)a + tb \in U \text{ for all } 0 \leq t \leq 1. \quad (2.70)$$

Before we state and prove the Inverse Function Theorem, we give the following definition.

Definition 2.13. Let U and V be open sets in \mathbb{R}^n and $f : U \rightarrow V$ a C^r map. The map f is a C^r *diffeomorphism* if it is bijective and $f^{-1} : V \rightarrow U$ is also C^r .

Inverse Function Theorem. Let U be an open set in \mathbb{R}^n , $f : U \rightarrow \mathbb{R}^n$ a C^r map, and $a \in U$. If $Df(a) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is bijective, then there exists a neighborhood U_1 of a in U and a neighborhood V of $f(a)$ in \mathbb{R}^n such that $F|_{U_1}$ is a C^r diffeomorphism of U_1 at V .

Proof. To prove this we need some elementary multi-variable calculus results, which we provide with the following lemmas.

Lemma 2.14. Let U be open in \mathbb{R}^n and $F : U \rightarrow \mathbb{R}^k$ be a C^1 mapping. Also assume that U is convex. Suppose that $|Df(a)| \leq c$ for all $A \in U$. Then, for all $x, y \in U$,

$$|f(x) - f(y)| \leq nc|x - y|. \quad (2.71)$$

Proof. Consider any $x, y \in U$. The Mean Value Theorem says that for every i there exists a point c on the line joining x to y such that

$$f_i(x) - f_i(y) = \sum_j \frac{\partial f_i}{\partial x_j}(d)(x_j - y_j). \quad (2.72)$$

It follows that

$$\begin{aligned} |f_i(x) - f_i(y)| &\leq \sum_j \left| \frac{\partial f_i}{\partial x_j}(d) \right| |x_j - y_j| \\ &\leq \sum_j c |x_j - y_j| \\ &\leq nc|x - y| \end{aligned} \quad (2.73)$$

This is true for each i , so $|f(x) - f(y)| \leq nc|x - y|$ □

Lemma 2.15. *Let U be open in \mathbb{R}^n and $f : U \rightarrow \mathbb{R}$ a \mathcal{C}^1 map. Suppose f takes a minimum value at some point $b \in U$. Then*

$$\frac{\partial f}{\partial x_i}(b) = 0, \quad i = 1, \dots, n. \quad (2.74)$$

Proof. We reduce to the one-variable result. Let $b = (b_1, \dots, b_n)$ and let $\phi(t) = f(b_1, \dots, b_{i-1}, t, b_{i+1}, \dots, b_n)$, which is \mathcal{C}^1 near b_1 and has a minimum at b_i . We know from one-variable calculus that this implies that $\frac{\partial \phi}{\partial t}(b_i) = 0$. □

In our proof of the Inverse Function Theorem, we want to show that f is locally a diffeomorphism at a . We will make the following simplifying assumptions:

$$a = 0, \quad f(a) = 0, \quad Df(0) = I \text{ (identity)}. \quad (2.75)$$

Then, we define a map $g : U \rightarrow \mathbb{R}^n$ by $g(x) = x - f(x)$, so that we obtain the further simplification

$$Dg(0) = Df(0) - I = 0. \quad (2.76)$$

Lemma 2.16. *Given $\epsilon > 0$, there exists $\delta > 0$ such that for any $x, y \in R_\delta$,*

$$|g(x) - g(y)| < \epsilon|x - y|. \quad (2.77)$$

Proof. The result that $Dg(0) = 0$ implies that there exists $\delta > 0$ such that for any $x \in R_\delta$, $|Dg(x)| \leq \epsilon/n$. Applying the first lemma, the proof is complete. □

Now, remember that $g(x) = x - f(x)$. Take any $x, y \in R_\delta$. Then

$$\begin{aligned} x - y &= x - f(x) + f(x) - f(y) + f(y) - y \\ &= g(x) - g(y) + f(x) - f(y). \end{aligned} \quad (2.78)$$

Using the Triangle Inequality we obtain

$$|x - y| \leq |g(x) - g(y)| + |f(x) - f(y)| \quad (2.79)$$

Using the previous lemma, we find that

$$(1 - \epsilon)|x - y| \leq |f(x) - f(y)|. \quad (2.80)$$

We choose δ such that $\epsilon > 1/2$, so that

$$|x - y| \leq 2|f(x) - f(y)|. \quad (2.81)$$

This proves that $f : R_\delta \rightarrow \mathbb{R}^n$ is one-to-one.

We also want to prove that f is onto. We have $Df(0) = I$, so $\det(\frac{\partial f_i}{\partial x_j}(0)) = 1$. We can choose δ such that for any $x \in R_\delta$,

$$\det \left(\frac{\partial f_i}{\partial x_j}(x) \right) > \frac{1}{2}. \quad (2.82)$$

Lemma 2.17. *If $y \in B_{\delta/4}$, then there exists a point $c \in R_\delta$ such that $f(c) = y$.*

Proof. Let $h : \bar{R}_\delta \rightarrow \mathbb{R}$ be a map defined by $h(x) = \|f(x) - y\|^2$. The domain \bar{R}_δ is compact, so h has a minimum at some point $c \in \bar{R}_\delta$.

Claim. *The point c is an interior point. That is, $c \in R_\delta$.*

Proof. For any $x \in \bar{R}_\delta$, $|x| = \delta$ implies that $|f(x) - f(0)| = |f(x)| \geq \delta/2$

$$\begin{aligned} \implies \|f(x)\| &\geq \frac{\delta}{2} \\ \implies \|f(x) - y\| &\geq \frac{\delta}{4}, \text{ when } x \in \text{Bd } R_\delta. \\ \implies h(x) &\geq \left(\frac{\delta}{4}\right)^2. \end{aligned} \quad (2.83)$$

At the origin, $h(0) = \|f(0) - y\|^2 = \|y\|^2 < (\delta/4)^2$, since $y \in B_{\delta/4}$. So, $h(0) \leq h$ on $\text{Bd } R_\delta$, which means that the minimum point c of h is in R_δ . This ends the proof of the claim. \square

Now that we know that the minimum point c occurs in the interior, we can apply the second lemma to h to obtain

$$\frac{\partial h}{\partial x_j}(c) = 0, \quad j = 1, \dots, n. \quad (2.84)$$

From the definition of h ,

$$h(x) = \sum_{i=1}^n (f_i(c) - y_i) \frac{\partial f_i}{\partial x_j}(c) = 0, \quad i = 1, \dots, n, \quad (2.85)$$

so

$$\frac{\partial h}{\partial x_j}(c) = 2 \sum_{i=1}^n (f_i(c) - y_i) \frac{\partial f_i}{\partial x_j}(c) = 0, \quad i = 1, \dots, n. \quad (2.86)$$

Note that

$$\det \left[\frac{\partial f_i}{\partial x_j}(c) \right] \neq 0, \quad (2.87)$$

so, by Cramer's Rule,

$$f_i(c) - y_i = 0, \quad i = 1, \dots, n. \quad (2.88)$$

□

Let $U_1 = R_\delta \sim f^{-1}(B_{\delta/4})$, where we have chosen $V = B_{\delta/4}$. We have shown that f is a bijective map.

Claim. *The map $f^{-1} : V \rightarrow U_1$ is continuous.*

Proof. Let $a, b \in V$, and define $x = f^{-1}(a)$ and $y = f^{-1}(b)$. Then $a = f(x)$ and $b = f(y)$.

$$|a - b| = |f(x) - f(y)| \geq \frac{\partial |x - y|}{\partial 2}, \quad (2.89)$$

so

$$|a - b| \geq \frac{1}{2} |f^{-1}(a) - f^{-1}(b)|. \quad (2.90)$$

This shows that f^{-1} is continuous on $V = B_{\delta/4}$. □

As a last item for today's lecture, we show the following:

Claim. *The map f^{-1} is differentiable at 0, and $Df^{-1}(0) = I$.*

Proof. Let $k \in \mathbb{R}^n - \{0\}$ and choose $k \neq 0$. We are trying to show that

$$\frac{f^{-1}(0 + k) - f^{-1}(0) - Df^{-1}(0)k}{|k|} \rightarrow 0 \text{ as } k \rightarrow 0. \quad (2.91)$$

We simplify

$$\frac{f^{-1}(0 + k) - f^{-1}(0) - Df^{-1}(0)k}{|k|} = \frac{f^{-1}(k) - k}{|k|}. \quad (2.92)$$

Define $h = f^{-1}(k)$ so that $k = f(h)$ and $|k| \leq 2|h|$. To show that

$$\frac{f^{-1}(k) - k}{|k|} \rightarrow 0 \text{ as } k \rightarrow 0, \quad (2.93)$$

it suffices to show that

$$\frac{f^{-1}(k) - k}{|h|} \rightarrow 0 \text{ as } h \rightarrow 0. \quad (2.94)$$

That is, it suffices to show that

$$\frac{h - f(h)}{|h|} \rightarrow 0 \text{ as } h \rightarrow 0. \quad (2.95)$$

But this is equal to

$$-\frac{f(h) - f(0) - Df(0)h}{|h|}, \quad (2.96)$$

which goes to zero as $h \rightarrow 0$ because f is differentiable at zero. \square

The proof of the Inverse Function Theorem continues in the next lecture.

Lecture 7

We continue our proof of the Inverse Function Theorem.

As before, we let U be an open set in \mathbb{R}^n , and we assume that $0 \in U$. We let $f : U \rightarrow \mathbb{R}^n$ be a \mathcal{C}^1 map, and we assume $f(0) = 0$ and that $Df(0) = I$. We summarize what we have proved so far in the following theorem.

Theorem 2.18. *There exists a neighborhood U_0 of 0 in U and a neighborhood V of 0 in \mathbb{R}^n such that*

1. f maps U_0 bijectively onto V
2. $f^{-1} : V \rightarrow U_0$ is continuous,
3. f^{-1} is differentiable at 0.

Now, we let U be an open set in \mathbb{R}^n , and we let $f : U \rightarrow \mathbb{R}^n$ be a \mathcal{C}^2 map, as before, but we return to our original assumptions that $a \in U$, $b = f(a)$, and $Df(a) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is bijective. We prove the following theorem.

Theorem 2.19. *There exists a neighborhood U_0 of a in U and a neighborhood V of b in \mathbb{R}^n such that*

1. f maps U_0 bijectively onto V
2. $f^{-1} : V \rightarrow U_0$ is continuous,
3. f^{-1} is differentiable at b .

Proof. The map $f : U \rightarrow \mathbb{R}^n$ maps a to b . Define $U' = U - a = \{x - a : x \in U\}$. Also define $f_1 : U' \rightarrow \mathbb{R}^n$ by $f_1(x) = f(x + a) - b$, so that $f_1(0) = 0$ and $Df_1(0) = Df(a)$ (using the Chain Rule).

Let $A = Df(a) = Df_1(0)$. We know that A is invertible.

Now, define $f_2 : U' \rightarrow \mathbb{R}^n$ by $f_2 = A^{-1}f_1$, so that $f_2(0) = 0$ and $Df_2(0) = I$. The results from last lecture show that the theorem at hand is true for f_2 . Because $f_1 = A \circ f_2$, the theorem is also true for f_1 . Finally, because $f(x) = f_1(x - a) + b$, the theorem is true for f . □

So, we have a bijective map $f : U_0 \rightarrow V$. Let us take $c \in U_0$ and look at the derivative

$$Df(c) \sim \left[\frac{\partial f_i}{\partial x_j}(c) \right] = J_f(c). \tag{2.97}$$

Note that

$$Df(c) \text{ is bijective} \iff \det \left[\frac{\partial f_i}{\partial x_j}(c) \right] \neq 0. \tag{2.98}$$

Because f is \mathcal{C}^1 , the functions $\frac{\partial f_i}{\partial x_j}$ are continuous on U_0 . If $\det J_f(a) \neq 0$, then $\det J_f(c) \neq 0$ for c close to a . We can shrink U_0 and V such that $\det J_f(c) \neq 0$ for

all $c \in U_0$, so for every $c \in U_0$, the map f^{-1} is differentiable at $f(c)$. That is, f^{-1} is differentiable at all points of V .

We have thus improved the previous theorem. We can replace the third point with

$$3. f^{-1} \text{ is differentiable at all points of } V. \quad (2.99)$$

Let $f^{-1} = g$, so that $g \circ f = \text{identity map}$. The Chain Rule is used to show the following. Suppose $p \in U_0$ and $q = f(p)$. Then $Dg(q) = Df(p)^{-1}$, so $J_g(q) = J_f(p)^{-1}$. That is, for all $x \in V$,

$$\left[\frac{\partial g_i}{\partial x_j}(x) \right] = \left[\frac{\partial f_i}{\partial x_j}(g(x)) \right]^{-1}. \quad (2.100)$$

The function f is \mathcal{C}^1 , so $\frac{\partial f_i}{\partial x_j}$ is continuous on U_0 . It also follows that g is continuous, so $\frac{\partial f_i}{\partial x_j}(g(x))$ is continuous on V .

Using Cramer's Rule, we conclude that the entries of matrix on the r.h.s. of Equation 2.100 are continuous functions on V . This shows that $\frac{\partial f_i}{\partial x_j}$ is continuous on V , which implies that g is a \mathcal{C}^1 map.

We leave as an exercise to show that $f \in \mathcal{C}^r$ implies that $g \in \mathcal{C}^r$ for all r . The proof is by induction.

This concludes the proof of the Inverse Function Theorem, signifying the end of this section of the course.

□

3 Integration

3.1 Riemann Integral of One Variable

We now begin to study the next main topic of this course: integrals. We begin our discussion of integrals with an 18.100 level review of integrals.

We begin by defining the Riemann integral (sometimes written in shorthand as the R. integral).

Let $[a, b] \subseteq \mathbb{R}$ be a closed interval in \mathbb{R} , and let P be a finite subset of $[a, b]$. Then P is a *partition* if $a, b \in P$ and if all of the elements t_1, \dots, t_N in P can be arranged such that $t_1 = a < t_2 < \dots < t_n = b$. We define $I_i = [t_i, t_{i+1}]$, which are called the subintervals of $[a, b]$ belonging to P .

Let $f : [a, b] \rightarrow \mathbb{R}$ be a bounded function, and let I_i be a subinterval belonging to P . Then we define

$$\begin{aligned} m_i &= \inf f : I_i \rightarrow \mathbb{R} \\ M_i &= \sup f : I_i \rightarrow \mathbb{R}, \end{aligned} \quad (3.1)$$

from which we define the *lower* and *upper* Riemann sums

$$\begin{aligned} L(f, P) &= \sum_i m_i \times (\text{length of } I_i) \\ U(f, P) &= \sum_i M_i \times (\text{length of } I_i), \end{aligned} \tag{3.2}$$

respectively.

Clearly,

$$L(f, P) \leq U(f, P). \tag{3.3}$$

Now, let P and P' be partitions.

Definition 3.1. The partition P is a *refinement* of P' if $P' \supset P$.

If P' is a refinement of P , then

$$\begin{aligned} L(f, P') &\geq L(f, P), \text{ and} \\ U(f, P') &\leq U(f, P). \end{aligned} \tag{3.4}$$

If P and P' are *any* partitions, then you can take $P'' = P \cup P'$, which is a refinement of both P and P' . So,

$$L(f, P) \leq L(f, P'') \leq U(f, P'') \leq U(f, P') \tag{3.5}$$

for any partitions P, P' . That is, the lower Riemann sum is always less than or equal to the upper Riemann sum, regardless of the partitions used.

Now we can define the Lower and Upper Riemann integrals

$$\begin{aligned} \int_{[a,b]} f &= \text{l.u.b. } \{L(f, P) | P \text{ a partition of } [a, b]\} \\ \overline{\int}_{[a,b]} f &= \text{g.l.b. } \{U(f, P) | P \text{ a partition of } [a, b]\} \end{aligned} \tag{3.6}$$

We can see from the above that

$$\underline{\int} f \leq \overline{\int} f. \tag{3.7}$$

Claim. If f is continuous, then

$$\underline{\int} f = \overline{\int} f. \tag{3.8}$$

Definition 3.2. For any bounded function $f : [a, b] \rightarrow \mathbb{R}$, the function f is (*Riemann*) *integrable* if

$$\int_{[a,b]} f = \overline{\int}_{[a,b]} f. \tag{3.9}$$

In the next lecture we will begin to generalize these notions to multiple variables.

Lecture 8

3.2 Riemann Integral of Several Variables

Last time we defined the Riemann integral for one variable, and today we generalize to many variables.

Definition 3.3. A *rectangle* is a subset Q of \mathbb{R}^n of the form

$$Q = [a_1, b_1] \times \cdots \times [a_n, b_n], \quad (3.10)$$

where $a_i, b_i \in \mathbb{R}$.

Note that $x = (x_1, \dots, x_n) \in Q \iff a_i \leq x_i \leq b_i$ for all i . The volume of the rectangle is

$$v(Q) = (b_1 - a_1) \cdots (b_n - a_n), \quad (3.11)$$

and the width of the rectangle is

$$\text{width}(Q) = \sup_i (b_i - a_i). \quad (3.12)$$

Recall (stated informally) that given $[a, b] \in \mathbb{R}$, a finite subset P of $[a, b]$ is a partition of $[a, b]$ if $a, b \in P$ and you can write $P = \{t_i : i = 1, \dots, N\}$, where $t_1 = a < t_2 < \dots < t_N = b$. An interval I belongs to P if and only if I is one of the intervals $[t_i, t_{i+1}]$.

Definition 3.4. A *partition* P of Q is an n -tuple (P_1, \dots, P_n) , where each P_i is a partition of $[a_i, b_i]$.

Definition 3.5. A rectangle $R = I_1 \times \cdots \times I_n$ *belongs to* P if for each i , the interval I_i belongs to P_i .

Let $f : Q \rightarrow \mathbb{R}$ be a bounded function, let P be a partition of Q , and let R be a rectangle belonging to P .

We define

$$\begin{aligned} m_R f &= \inf_R f = \text{g.l.b. } \{f(x) : x \in R\} \\ M_R f &= \sup_R f = \text{l.u.b. } \{f(x) : x \in R\}, \end{aligned} \quad (3.13)$$

from which we define the lower and upper Riemann sums,

$$\begin{aligned} L(f, P) &= \sum_R m_R(f) v(R) \\ U(f, P) &= \sum_R M_R(f) v(R). \end{aligned} \quad (3.14)$$

It is evident that

$$L(f, P) \leq U(f, P). \quad (3.15)$$

Now, we will take a sequence of partitions that get finer and finer, and we will define the integral to be the limit.

Let $P = (P_1, \dots, P_n)$ and $P' = (P'_1, \dots, P'_n)$ be partitions of Q . We say that P' refines P if $P'_i \supset P_i$ for each i .

Claim. *If P' refines P , then*

$$L(f, P') \geq L(f, P). \quad (3.16)$$

Proof. We let $P_j = P'_j$ for $j \neq i$, and we let $P'_i = P_i \cup \{a\}$, where $a \in [a_i, b_i]$. We can create any refinement by multiple applications of this basic refinement. If R is a rectangle belonging to P , then either

1. R belongs to P' , or
2. $R = R' \cup R''$, where R', R'' belong to P' .

In the first case, the contribution of R to $L(f, P')$ equals the contribution of R to $L(f, P)$, so the claim holds.

In the second case,

$$m_R v(R) = m_R (v(R') + v(R'')) \quad (3.17)$$

and

$$m_r = \inf_R f \leq \inf_{R'} f, \inf_{R''} f. \quad (3.18)$$

So,

$$m_R \leq m_{R'}, m_{R''} \quad (3.19)$$

Taken altogether, this shows that

$$m_R v(R) \leq m_{R'} v(R') + m_{R''} v(R'') \quad (3.20)$$

Thus, R' and R'' belong to P' . □

Claim. *If P' refines P , then*

$$U(f, P') \leq U(f, P) \quad (3.21)$$

The proof is very similar to the previous proof. Combining the above two claims, we obtain the corollary

Corollary 2. *If P and P' are partitions, then*

$$U(f, P') \geq L(f, P) \quad (3.22)$$

Proof. Define $P'' = (P_1 \cup P'_1, \dots, P_n \cup P'_n)$. So, P'' refines P and P' . We have shown that

$$\begin{aligned} U(f, P'') &\leq U(f, P) \\ L(f, P') &\leq L(f, P'') \\ L(f, P'') &\leq U(f, P''). \end{aligned} \tag{3.23}$$

Together, these show that

$$U(f, P) \geq L(f, P'). \tag{3.24}$$

□

With this result in mind, we define the lower and upper Riemann integrals:

$$\begin{aligned} \int_{\underline{Q}} f &= \sup_P L(f, P) \\ \int_{\overline{Q}} f &= \inf_P U(f, P). \end{aligned} \tag{3.25}$$

Clearly, we have

$$\int_{\underline{Q}} f \leq \int_{\overline{Q}} f, \tag{3.26}$$

Finally, we define Riemann integrable.

Definition 3.6. A function f is *Riemann integrable over Q* if the lower and upper Riemann integrals coincide (are equal).

3.3 Conditions for Integrability

Our next problem is to determine under what conditions a function is (Riemann) integrable.

Let's look at a trivial case:

Claim. Let $F : Q \rightarrow \mathbb{R}$ be the constant function $f(x) = c$. Then f is *R. integrable over Q* , and

$$\int_Q c = cv(Q). \tag{3.27}$$

Proof. Let P be a partition, and let R be a rectangle belonging to P . We see that $m_R(f) = M_R(f) = c$, so

$$\begin{aligned} U(f, P) &= \sum_R M_R(f)v(R) = c \sum_R v(R) \\ &= cv(Q). \end{aligned} \tag{3.28}$$

Similarly,

$$L(f, P) = cv(Q). \quad (3.29)$$

□

Corollary 3. Let Q be a rectangle, and let $\{Q_i : i = 1, \dots, N\}$ be a collection of rectangles covering Q . Then

$$v(Q) \leq \sum v(Q_i). \quad (3.30)$$

Theorem 3.7. If $f : Q \rightarrow \mathbb{R}$ is continuous, then f is R . integrable over Q .

Proof. We begin with a definition

Definition 3.8. Given a partition P of Q , we define

$$\text{mesh width}(P) = \sup_R \text{width}(R). \quad (3.31)$$

Remember that

$$Q \text{ compact} \implies f : Q \rightarrow \mathbb{R} \text{ is uniformly continuous.} \quad (3.32)$$

That is, given $\epsilon > 0$, there exists $\delta > 0$ such that if $x, y \in Q$ and $|x - y| < \delta$, then $|f(x) - f(y)| < \epsilon$.

Choose a partition P of Q with mesh width less than δ . Then, for every rectangle R belonging to P and for every $x, y \in R$, we have $|x - y| < \delta$. By uniform continuity we have, $M_R(f) - m_R(f) \leq \epsilon$, which is used to show that

$$\begin{aligned} U(f, P) - L(f, P) &= \sum_R (M_R(f) - m_R(f))v(R) \\ &\leq \epsilon \sum v(R) \\ &\leq \epsilon v(Q). \end{aligned} \quad (3.33)$$

We can take $\epsilon \rightarrow 0$, so

$$\sup_P L(f, P) = \inf_P U(f, P), \quad (3.34)$$

which shows that f is integrable. □

We have shown that continuity is sufficient for integrability. However, continuity is clearly not necessary. What is the general condition for integrability? To state the answer, we need the notion of *measure zero*.

Definition 3.9. Suppose $A \subseteq \mathbb{R}^n$. The set A is of *measure zero* if for every $\epsilon > 0$, there exists a countable covering of A by rectangles Q_1, Q_2, Q_3, \dots such that $\sum_i v(Q_i) < \epsilon$.

Theorem 3.10. *Let $f : Q \rightarrow \mathbb{R}$ be a bounded function, and let $A \subseteq Q$ be the set of points where f is not continuous. Then f is R. integrable if and only if A is of measure zero.*

Before we prove this, we make some observations about sets of measure zero:

1. Let $A, B \subseteq \mathbb{R}^n$ and suppose $B \subset A$. If A is of measure zero, then B is also of measure zero.
2. Let $A_i \subseteq \mathbb{R}^n$ for $i = 1, 2, 3, \dots$, and suppose the A_i 's are of measure zero. Then $\cup A_i$ is also of measure zero.
3. Rectangles are *not* of measure zero.

We prove the second observation:

For any $\epsilon > 0$, choose coverings $Q_{i,1}, Q_{i,2}, \dots$ of A_i such that each covering has total volume less than $\epsilon/2^i$. Then $\{Q_{i,j}\}$ is a countable covering of $\cup A_i$ of total volume

$$\sum_{i=1}^{\infty} \frac{\epsilon}{2^i} = \epsilon. \quad (3.35)$$

Lecture 9

We quickly review the definition of measure zero.

A set $A \subseteq \mathbb{R}^n$ is of *measure zero* if for every $\epsilon > 0$, there exists a covering of A by rectangles Q_1, Q_2, Q_3, \dots such that the total volume $\sum v(Q_i) < \epsilon$.

Remark. In this definition we can replace “rectangles” by “open rectangles.” To see this, given any $\epsilon > 0$ let Q_1, Q_2, \dots be a cover of A with volume less than $\epsilon/2$. Next, choose Q'_i to be rectangles such that $\text{Int } Q'_i \supset Q_i$ and $v(Q'_i) < 2v(Q_i)$. Then $\text{Int } Q'_1, \text{Int } Q'_2, \dots$ cover A and have total volume less than ϵ .

We also review the three properties of measure zero that we mentioned last time, and we prove the third.

1. Let $A, B \subseteq \mathbb{R}^n$ and suppose $B \subset A$. If A is of measure zero, then B is also of measure zero.
2. Let $A_i \subseteq \mathbb{R}^n$ for $i = 1, 2, 3, \dots$, and suppose the A_i 's are of measure zero. Then $\cup A_i$ is also of measure zero.
3. Rectangles are *not* of measure zero.

We prove the third property:

Claim. *If Q is a rectangle, then Q is not of measure zero.*

Proof. Choose $\epsilon < v(Q)$. Suppose Q_1, Q_2, \dots are rectangles such that the total volume is less than ϵ and such that $\text{Int } Q_1, \text{Int } Q_2, \dots$ cover Q .

The set Q is compact, so the H-B Theorem implies that the collection of sets $\text{Int } Q_1, \dots, \text{Int } Q_N$ cover Q for N sufficiently large. So,

$$Q \subseteq \bigcup_{i=1}^N Q_i, \tag{3.36}$$

which implies that

$$v(Q) \leq \sum_{i=1}^N v(Q_i) < \epsilon < v(Q), \tag{3.37}$$

which is a contradiction. □

We then have the following simple result.

Claim. *If $\text{Int } A$ is non-empty, then A is not of measure zero.*

Proof. Consider any $p \in \text{Int } A$. There exists a $\delta > 0$ such that $U(p, \delta) = \{x : |x - p| < \delta\}$ is contained in A . Then let $Q = \{x : |x - p| \leq \delta\}$. It follows that if A is of measure zero, then Q is of measure zero, by the first property. We know that Q is not of measure zero by the third property. □

We restate the necessary and sufficient condition for R. integrability from last time, and we now prove the theorem.

Theorem 3.11. *Let Q be a rectangle and $f : Q \rightarrow \mathbb{R}$ be a bounded function. Let D be the set of points in Q where f is not continuous. Then f is R. integrable if and only if D is of measure zero.*

Proof. First we show that

$$D \text{ is of measure zero} \implies f \text{ is R. integrable} \quad (3.38)$$

Lemma 3.12. *Let $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$, and let $Q^\alpha, \alpha = 1, \dots, N$, be a covering of Q by rectangles. Then there exists a partition P of Q such that every rectangle R belonging to P is contained in Q^α for some α .*

Proof. Write out $Q^\alpha = I_1^\alpha \times \cdots \times I_n^\alpha$, and let

$$P_j = \left(\bigcup_{\alpha} \text{Endpoints of } I_j^\alpha \right) \cap [a_j, b_j] \cup \{a_j, b_j\}. \quad (3.39)$$

One can show that P_j is a partition of $[a_j, b_j]$, and $P = (P_1, \dots, P_n)$ is a partition of Q with the above properties. \square

Let $f : Q \rightarrow \mathbb{R}$ be a bounded function, and let D be the set of points at which f is discontinuous. Assume that D is of measure zero. We want to show that f is R. integrable.

Let $\epsilon > 0$, and let $Q'_i, i = 1, 2, 3, \dots$ be a collection of rectangles of total volume less than ϵ such that $\text{Int } Q'_1, Q'_2, \dots$ cover D .

If $p \in Q - D$, we know that f is continuous at p . So, there exists a rectangle Q_p with $p \in \text{Int } Q_p$ and $|f(x) - f(p)| < \epsilon/2$ for all $x \in Q_p$ (for example, $Q_p = \{x \mid |x - p| \leq \delta\}$ for some δ). Given any $x, y \in Q_p$, we find that $|f(x) - f(y)| < \epsilon$.

The rectangles $\text{Int } Q_p, p \in Q - D$ along with the rectangles $\text{Int } Q'_i, i = 1, 2, \dots$ cover Q . The set Q is compact, so the H-B Theorem implies that there exists a finite open subcover:

$$Q_i \equiv \text{Int } Q_{p_i}, i = 1, \dots, \ell; \quad \text{Int } Q'_j, j = 1, \dots, \ell. \quad (3.40)$$

Using the lemma, there exists a partition P of Q such that every rectangle belonging to P is contained in a Q_i or a Q'_j .

We now show that f is R. integrable.

$$\begin{aligned} U(f, P) - L(f, P) &= \sum_R (M_R(f) - m_R(f))v(R) \\ &\quad + \sum_{R'} (M_{R'}(f) - m_{R'}(f))v(R'), \end{aligned} \quad (3.41)$$

where each R in the first sum belongs to a Q_i , and each R' in the second sum belongs to a Q'_j .

We look at the first sum. If $x, y \in R \subseteq Q_i$, then $|f(x) - f(y)| \leq \epsilon$. So, $M_R(f) - m_R(f) \leq \epsilon$. It follows that

$$\begin{aligned} \sum_R (M_R(f) - m_R(f))v(R) &\leq \epsilon \sum_R v(R) \\ &\leq \epsilon v(Q). \end{aligned} \tag{3.42}$$

We now look at the second sum. The function $f : Q \rightarrow \mathbb{R}$ is bounded, so there exists a number c such that $-c \leq f(x) \leq c$ for all $x \in Q$. Then, $M_{R'}(f) - m_{R'}(f) \leq 2c$ so

$$\begin{aligned} \sum_{R'} (M_{R'}(f) - m_{R'}(f))v(R') &\leq 2c \sum_{R'} v(R') \\ &= 2c \sum_{i=1}^{\ell} \sum_{R' \subseteq Q'_i} v(R') \\ &\leq 2c \sum_i v(Q'_i) \\ &\leq 2c\epsilon. \end{aligned} \tag{3.43}$$

Substituting back into Equation 3.41, we get

$$U(f, P) - L(f, P) \leq \epsilon(v(Q) + 2c). \tag{3.44}$$

So,

$$\overline{\int}_Q f - \underline{\int}_Q f \leq \epsilon(v(Q) + 2c), \tag{3.45}$$

because

$$U(f, P) \geq \overline{\int}_Q f \text{ and } L(f, P) \leq \underline{\int}_Q f. \tag{3.46}$$

Letting ϵ go to zero, we conclude that

$$\overline{\int}_Q f = \underline{\int}_Q f, \tag{3.47}$$

which shows that f is Riemann integrable.

This concludes the proof in one direction. We do not prove the other direction. \square

Corollary 4. *Suppose $f : Q \rightarrow \mathbb{R}$ is R. integrable and that $f \geq 0$ everywhere. If $\int_Q f = 0$, then $f = 0$ except on a set of measure zero.*

Proof. Let D be the set of points where f is discontinuous. The function f is R. integrable, so D is of measure zero.

If $p \in Q - D$, then $f(p) = 0$. To see this, suppose that $f(p) = \delta > 0$. The function f is continuous at p , so there exists a rectangle R_0 centered at p such that $f \geq n\delta/2$ on R_0 . Choose a partition P such that R_0 is a rectangle belonging to P . On any rectangle R belonging to P , $f \geq 0$, so $m_R(f) \geq 0$. This shows that

$$\begin{aligned} L(f, P) &= m_{R_0}(f)v(R_0) + \sum_{R \neq R_0} m_R(f)v(R) \\ &\geq \frac{\delta}{2}v(R_0) + 0. \end{aligned} \tag{3.48}$$

But we assumed that $\int_Q f = 0$, so we have reached a contradiction. So $f = 0$ at all points $p \in Q - D$. \square

Lecture 10

We begin today's lecture with a simple claim.

Claim. Let $Q \subseteq \mathbb{R}^n$ be a rectangle and $f, g : Q \rightarrow \mathbb{R}$ be bounded functions such that $f \leq g$. Then

$$\int_{\underline{Q}} f \leq \int_{\underline{Q}} g. \quad (3.49)$$

Proof. Let P be a partition of Q , and let R be a rectangle belonging to P . Clearly, $m_R(f) \leq m_R(g)$, so

$$L(f, P) = \sum_R m_R(f)v(R) \quad (3.50)$$

$$L(g, P) = \sum_R m_R(g)v(R) \quad (3.51)$$

$$\implies L(f, P) \leq L(g, P) \leq \int_{\underline{Q}} g, \quad (3.52)$$

for all partitions P . The lower integral

$$\int_{\underline{Q}} f \quad (3.53)$$

is the l.u.b. of $L(f, P)$, so

$$\int_{\underline{Q}} f \leq \int_{\underline{Q}} g. \quad (3.54)$$

□

Similarly,

$$\overline{\int}_Q f \leq \overline{\int}_Q g. \quad (3.55)$$

It follows that if $f \leq g$, then

$$\int_Q f \leq \int_Q g. \quad (3.56)$$

This is the *monotonicity* property of the R. integral.

3.4 Fubini Theorem

In one-dimensional calculus, when we have a continuous function $f : [a, b] \rightarrow \mathbb{R}$, then we can calculate the R. integral

$$\int_a^b f(x)dx = F(b) - F(a), \quad (3.57)$$

where F is the anti-derivative of f .

When we integrate a continuous function $f : Q \rightarrow \mathbb{R}$ over a two-dimensional region, say $Q = [a_1, b_1] \times [a_2, b_2]$, we can calculate the R. integral

$$\int_Q f = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y) dx dy = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x, y) dx dy \right) \quad (3.58)$$

That is, we can break up Q into components and integrate separately over those components. We make this more precise in the following Fubini Theorem.

First, we define some notation that will be used.

Let $n = k + \ell$ so that $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^\ell$. Let $c = (c_1, \dots, c_n) \in \mathbb{R}^n$. We can write $c = (a, b)$, where $a = (c_1, \dots, c_k) \in \mathbb{R}^k$ and $b = (c_{k+1}, \dots, c_{k+\ell}) \in \mathbb{R}^\ell$. Similarly, let $Q = I_1 \times \dots \times I_n$ be a rectangle in \mathbb{R}^n . Then we can write $Q = A \times B$, where $A = I_1 \times \dots \times I_k \in \mathbb{R}^k$ and $B = I_{k+1} \times \dots \times I_{k+\ell} \in \mathbb{R}^\ell$. Along the same lines, we can write a partition $P = (P_1, \dots, P_n)$ as $P = (P_A, P_B)$, where $P_A = (P_1, \dots, P_k)$ and $P_B = (P_{k+1}, \dots, P_{k+\ell})$.

Fubini Theorem. *Let $f : Q \rightarrow \mathbb{R}$ 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 $f_x : B \rightarrow \mathbb{R}$ by $f_x(y) = f(x, y)$. Since this function is bounded, we can define new functions $g, h : A \rightarrow \mathbb{R}$ by*

$$g(x) = \int_{\underline{B}} f_x, \quad (3.59)$$

$$h(x) = \int_{\overline{B}} f_x. \quad (3.60)$$

Note that $g \leq h$. The Fubini Theorem concludes the following: If f is integrable over Q , then g and h are integrable over A and

$$\int_A g = \int_A h = \int_Q f. \quad (3.61)$$

Proof. Let $P = (P_A, P_B)$ be a partition of Q , and let $R = R_A \times R_B$ be a rectangle belonging to P (so R_A belongs to P_A and R_B belongs to P_B). Fix $x_0 \in A$.

First, we claim that

$$m_{R_A \times R_B}(f) \leq m_{R_B}(f_{x_0}), \quad (3.62)$$

the proof of which is straightforward.

Next,

$$\begin{aligned} \sum_{R_B} m_{R_A \times R_B}(f) v(R_B) &\leq \sum_{R_B} m_{R_B}(f_{x_0}) v(R_B) \\ &= L(f_{x_0}, P_B) \\ &\leq \int_{\underline{B}} f_{x_0} = g(x_0). \end{aligned} \quad (3.63)$$

So,

$$\sum_{R_B} m_{R_A \times R_B}(f)v(R_B) \leq g(x_0) \quad (3.64)$$

for all $x_0 \in R_A$. The above equation must hold for the infimum of the r.h.s, so

$$\sum_{R_B} m_{R_A \times R_B}(f)v(R_B) \leq m_{R_A}(g). \quad (3.65)$$

Observe that $v(R_A \times R_B) = v(R_A)v(R_B)$, so

$$\begin{aligned} L(f, P) &= \sum_{R_A \times R_B} m_{R_A \times R_B}(f)v(R_A \times R_B) \\ &\leq \sum_{R_A} m_{R_A}(g)v(R_A) \\ &\leq \int_{\underline{A}} g. \end{aligned} \quad (3.66)$$

We have just shown that for any partition $P = (P_A, P_B)$,

$$L(f, P) \leq L(g, P_A) \leq \int_{\underline{A}} g, \quad (3.67)$$

so

$$\int_{\underline{Q}} f \leq \int_{\underline{A}} g. \quad (3.68)$$

By a similar argument, we can show that

$$\int_A h \leq \int_Q f. \quad (3.69)$$

Summarizing, we have shown that

$$\int_{\underline{Q}} f \leq \int_{\underline{A}} g \leq \int_A h \leq \int_Q f, \quad (3.70)$$

where we used monotonicity for the middle inequality. Since f is R. integrable,

$$\int_{\underline{Q}} f = \int_Q f, \quad (3.71)$$

so all of the inequalities are in fact equalities. \square

Remark. Suppose that for every $x \in A$, that $f_x : B \rightarrow \mathbb{R}$ is R. integrable. That's the same as saying $g(x) = h(x)$. Then

$$\begin{aligned} \int_A \left(\int_B f_x \right) &= \int_A dx \left(\int_B f(x, y) dy \right) \\ &= \int_{A \times B} f(x, y) dx dy, \end{aligned} \tag{3.72}$$

using standard notation from calculus.

Remark. In particular, if f is continuous, then f_x is continuous. Hence, the above remark holds for all continuous functions.

3.5 Properties of Riemann Integrals

We now prove some standard calculus results.

Theorem 3.13. *Let $Q \subseteq \mathbb{R}^n$ be a rectangle, and let $f, g : Q \rightarrow \mathbb{R}$ be R. integrable functions. Then, for all $a, b \in \mathbb{R}$, the function $af + bg$ is R. integrable and*

$$\int_Q af + bg = a \int_Q f + b \int_Q g. \tag{3.73}$$

Proof. Let's first assume that $a, b \leq 0$. Let P be a partition of Q and R a rectangle belonging to P . Then

$$am_R(f) + bm_R(g) \leq m_R(af + bg), \tag{3.74}$$

so

$$\begin{aligned} aL(f, P) + bL(g, P) &\leq L(af + bg, P) \\ &\leq \int_{\underline{Q}} af + bg. \end{aligned} \tag{3.75}$$

Claim. *For any pair of partitions P' and P'' ,*

$$aL(f, P') + bL(g, P'') \leq \int_{\underline{Q}} af + bg. \tag{3.76}$$

To see that the claim is true, take P to be a refinement of P' and P'' , and apply Equation 3.75. Thus,

$$a \int_{\underline{Q}} f + b \int_{\underline{Q}} g \leq \int_{\underline{Q}} af + bg. \tag{3.77}$$

Similarly, we can show that

$$\int_{\overline{Q}} af + bg \leq a \int_{\overline{Q}} f + b \int_{\overline{Q}} g. \tag{3.78}$$

Since f and g are R. integrable, we know that

$$\overline{\int}_Q f = \underline{\int}_Q f, \quad \overline{\int}_Q g = \underline{\int}_Q g. \quad (3.79)$$

These equalities show that the previous inequalities were in fact equalities, so

$$\int_Q af + bg = a \int_Q f + b \int_Q g. \quad (3.80)$$

However, remember that we assumed that $a, b \geq 0$. To deal with the case of arbitrary a, b , it suffices to check what happens when we change the sign of a or b .

Claim.

$$\int_Q -f = - \int_Q f. \quad (3.81)$$

Proof Hint. Let P be any partition of Q . Then $L(f, P) = -U(-f, P)$. □

You should check this claim, and then use it to complete the proof. □

Lecture 11

We review some basic properties of the Riemann integral.

Let $Q \subseteq \mathbb{R}^n$ be a rectangle, and let $f, g : Q \rightarrow \mathbb{R}$ be bounded functions. Assume that f, g are R. integrable. We have the following properties of R. integrals:

- Linearity: $a, b \in \mathbb{R} \implies af + bg$ is R. integrable and

$$\int_Q af + bg = a \int_Q f + b \int_Q g. \quad (3.82)$$

- Monotonicity: If $f \leq g$, then

$$\int_Q f \leq \int_Q g. \quad (3.83)$$

- Maximality Property: Let $h : Q \rightarrow \mathbb{R}$ be a function defined by $h(x) = \max(f(x), g(x))$.

Theorem 3.14. *The function h is R. integrable and*

$$\int_Q h \geq \max\left(\int_Q f, \int_Q g\right). \quad (3.84)$$

Proof. We need the following lemma.

Lemma 3.15. *If f and g are continuous at some point $x_0 \in Q$, then h is continuous at x_0 .*

Proof. We consider the case $f(x_0) = g(x_0) = h(x_0) = r$. The functions f and g are continuous at x_0 if and only if for every $\epsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| < \epsilon$ and $|g(x) - g(x_0)| < \epsilon$ whenever $|x - x_0| < \delta$.

Substitute in $f(x_0) = g(x_0) = r$. The value of $h(x)$ is either $f(x)$ or $g(x)$, so $|h(x) - r| < \epsilon$ for $|x - x_0| < \delta$. That is $|h(x) - h(x_0)| < \epsilon$ for $|x - x_0| < \delta$, so h is continuous at x_0 .

The proofs of the other cases are left to the student.

□

We defined $h = \max(f, g)$. The lemma tells us that h is integrable.

Define E, F , and G as follows:

$$E = \text{Set of points in } Q \text{ where } f \text{ is discontinuous,} \quad (3.85)$$

$$F = \text{Set of points in } Q \text{ where } g \text{ is discontinuous,} \quad (3.86)$$

$$G = \text{Set of points in } Q \text{ where } h \text{ is discontinuous.} \quad (3.87)$$

The functions f, g are integrable over Q if and only if E, F are of measure zero. The lemma shows that $G \subseteq E \cup F$, so h is integrable over Q . To finish the proof, we notice that

$$h = \max(f, g) \geq f, g. \quad (3.88)$$

Then, by monotonicity,

$$\int_Q h \geq \max\left(\int_Q f, \int_Q g\right). \quad (3.89)$$

□

Remark. Let $k = \min(f, g)$. Then $k = -\max(-f, -g)$. So, the maximality property also implies that k is integrable and

$$\int_Q k \leq \min\left(\int_Q f, \int_Q g\right). \quad (3.90)$$

A useful trick for when dealing with functions is to change the sign. The preceding remark and the following are examples where such a trick is useful.

Let $f : Q \rightarrow \mathbb{R}$ be a R. integrable function. Define

$$f_+ = \max(f, 0), \quad f_- = \max(-f, 0). \quad (3.91)$$

Both of these functions are R. integrable and non-negative: $f_+, f_- \geq 0$. Also note that $f = f_+ - f_-$. This decomposition is a trick we will use over and over again.

Also note that $|f| = f_+ + f_-$, so $|f|$ is R. integrable. By monotonicity,

$$\begin{aligned} \int_Q |f| &= \int_Q f_+ + \int_Q f_- \\ &\geq \int_Q f_+ - \int_Q f_- \\ &= \int_Q f. \end{aligned} \quad (3.92)$$

By replacing f by $-f$, we obtain

$$\begin{aligned} \int_Q |f| &\geq \int_Q -f \\ &= -\int_Q f. \end{aligned} \quad (3.93)$$

Combining these results, we arrive at the following claim

Claim.

$$\int_Q |f| \geq \left| \int_Q f \right| \quad (3.94)$$

Proof. The proof is above. □

3.6 Integration Over More General Regions

So far we've been defining integrals over rectangles. Let us now generalize to other sets.

Let S be a bounded set in \mathbb{R}^n , and let $f : S \rightarrow \mathbb{R}$ be a bounded function. Let $f_S : \mathbb{R}^n \rightarrow \mathbb{R}$ be the map defined by

$$f_S(x) = \begin{cases} f(x) & \text{if } x \in S, \\ 0 & \text{if } x \notin S. \end{cases} \quad (3.95)$$

Let Q be a rectangle such that $\text{Int } Q \supset \bar{S}$.

Definition 3.16. The map f is *Riemann integrable over S* if f_S is Riemann integrable over Q . Additionally,

$$\int_S f = \int_Q f_S. \quad (3.96)$$

One has to check that this definition does not depend on the choice of Q , but we do not check this here.

Claim. Let S be a bounded set in \mathbb{R}^n , and let $f, g : S \rightarrow \mathbb{R}$ be bounded functions. Assume that f, g are *R. integrable over S* . Then the following properties hold:

- *Linearity:* If $a, b \in \mathbb{R}$, then $af + bg$ is *R. integrable over S* , and

$$\int_S af + bg = a \int_S f + b \int_S g. \quad (3.97)$$

- *Monotonicity:* If $f \leq g$, then

$$\int_S f \leq \int_S g. \quad (3.98)$$

- *Maximality:* If $h = \max(f, g)$ (over the domain S), then h is *R. integrable over S* , and

$$\int_S h \geq \max\left(\int_S f, \int_S g\right). \quad (3.99)$$

Proof. The proofs are easy, and we outline them here.

- *Linearity:* Note that $af_S + bg_S = (af + bg)_S$, so

$$\begin{aligned} \int_S af + bg &= \int_Q (af + bg)_S \\ &= a \int_Q f_S + b \int_Q g_S \\ &= a \int_S f + b \int_S g. \end{aligned} \quad (3.100)$$

- Monotonicity: Use $f \leq g \implies f_S \leq g_S$.
- Maximality: Use $h = \max(f, g) \implies h_S = \max(f_S, g_S)$.

□

Let's look at some nice set theoretic properties of the Riemann integral.

Claim. Let S, T be bounded subsets of \mathbb{R}^n with $T \subseteq S$. Let $f : S \rightarrow \mathbb{R}$ be bounded and non-negative. Suppose that f is R. integrable over both S and T . Then

$$\int_T f \leq \int_S f. \quad (3.101)$$

Proof. Clearly, $f_T \leq f_S$. Let Q be a rectangle with $\bar{S} \supseteq \text{Int } Q$. Then

$$\int_Q f_T \leq \int_Q f_S. \quad (3.102)$$

□

Claim. Let S_1, S_2 be bounded subsets of \mathbb{R}^n , and let $f : S_1 \cup S_2 \rightarrow \mathbb{R}$ be a bounded function. Suppose that f is R. integrable over both S_1 and S_2 . Then f is R. integrable over $S_1 \cap S_2$ and over $S_1 \cup S_2$, and

$$\int_{S_1 \cup S_2} f = \int_{S_1} f + \int_{S_2} f - \int_{S_1 \cap S_2} f. \quad (3.103)$$

Proof. Use the following trick. Notice that

$$f_{S_1 \cup S_2} = \max(f_{S_1}, f_{S_2}), \quad (3.104)$$

$$f_{S_1 \cap S_2} = \min(f_{S_1}, f_{S_2}). \quad (3.105)$$

Now, choose Q such that

$$\text{Int } Q \supset \overline{S_1 \cup S_2}, \quad (3.106)$$

so $f_{S_1 \cup S_2}$ and $f_{S_1 \cap S_2}$ are integrable over Q .

Note the identity

$$f_{S_1 \cup S_2} = f_{S_1} + f_{S_2} - f_{S_1 \cap S_2}. \quad (3.107)$$

So,

$$\int_Q f_{S_1 \cup S_2} = \int_Q f_{S_1} + \int_Q f_{S_2} - \int_Q f_{S_1 \cap S_2}, \quad (3.108)$$

from which it follows that

$$\int_{S_1 \cup S_2} f = \int_{S_1} f + \int_{S_2} f - \int_{S_1 \cap S_2} f. \quad (3.109)$$

□

Lecture 12

So far, we have been studying only the Riemann integral. However, there is also the Lebesgue integral. These are the two basic integral theories. The Riemann integral is very intuitive and is usually adequate for problems that usually come up. The Lebesgue integral is not as intuitive, but it can handle more general problems. We do not encounter these problems in geometry or physics, but we would in probability and statistics. You can learn more about Lebesgue integrals by taking Fourier Analysis (18.103) or Measure and Integration (18.125). We do not study the Lebesgue integral.

Let S be a bounded subset of \mathbb{R}^n .

Theorem 3.17. *If the boundary of S is of measure zero, then the constant function 1 is R . integrable over S . The converse is also true.*

Proof. Let Q be a rectangle such that $\text{Int } Q \supset \bar{S}$. Define

$$1_S(x) = \begin{cases} 1 & \text{if } x \in S, \\ 0 & \text{if } x \notin S. \end{cases} \quad (3.110)$$

The constant function 1 is integrable over S if and only if the function 1_S is integrable over Q . The function 1_S is integrable over Q if the set of points D in Q where 1_S is discontinuous is of measure zero. If so, then

$$\int_Q 1_S = \int_S 1. \quad (3.111)$$

Let $x \in Q$.

1. If $x \in \text{Int } S$, then $1_S = 1$ in a neighborhood of x , so 1_S is continuous at x .
2. If $x \in \text{Ext } S$, then $1_S = 0$ in a neighborhood of x , so 1_S is continuous at x .
3. If $x \in \text{Bd } S$, then in every neighborhood U of x there exists points in $\text{Ext } S$ where $1_S = 0$ and points in $\text{Int } S$ where $1_S = 1$. So, 1_S is discontinuous at x .

Thus, D is the boundary of S , $D = \text{Bd } S$. Therefore, the function 1_S is integrable if and only if $\text{Bd } S$ is of measure zero. \square

3.7 Improper Integrals

Definition 3.18. The set S is *rectifiable* if the boundary of S is of measure zero. If S is rectifiable, then

$$v(S) = \int_S 1. \quad (3.112)$$

Let us look at the properties of $v(S)$:

1. Monotonicity: If S_1 and S_2 are rectifiable and $S_1 \subseteq S_2$, then $v(S_1) \leq v(S_2)$.
2. Linearity: If S_1, S_2 are rectifiable, then $S_1 \cup S_2$ and $S_1 \cap S_2$ are rectifiable, and

$$v(S_1 \cup S_2) = v(S_1) + v(S_2) - v(S_1 \cap S_2). \quad (3.113)$$
3. If S is rectifiable, then $v(S) = 0$ if and only if S is of measure zero.
4. Let $A = \text{Int } S$. If S is rectifiable, then A is rectifiable, and $v(S) = v(A)$.

The first two properties above are special cases of the theorems that we proved last lecture:

1.

$$\int_{S_1} 1 \leq \int_{S_2} 1 \quad \text{if } S_1 \subseteq S_2. \quad (3.114)$$

2.

$$\int_{S_1 \cup S_2} 1 = \int_{S_1} 1 + \int_{S_2} 1 - \int_{S_1 \cap S_2} 1. \quad (3.115)$$

To see the the third and fourth properties are true, we use some previous results. Let Q be a rectangle, and let $f : Q \rightarrow \mathbb{R}$ be R. integrable. We proved the following two theorems:

Theorem A. *If $f \geq 0$ and $\int_Q f = 0$, then $f = 0$ except on a set of measure zero.*

Theorem B. *If $f = 0$ except on a set of measure zero, then $\int_S f = 0$.*

Property 3. above is a consequence of Theorem A with $f = 1_S$.

Property 4. above is a consequence of Theorem B with $f = 1_S - 1_A$.

We are still lacking some simple criteria for a bounded set to be integrable. Let us now work on that.

Let S be a bounded set, and let $f : S \rightarrow \mathbb{R}$ be a bounded function. We want simple criteria on S and f such that f to be integrable over S .

Theorem 3.19. *If S is rectifiable and $f : S \rightarrow \mathbb{R}$ is bounded and continuous, then f is R. integrable over S .*

Proof. Let Q be a rectangle such that $\text{Int } Q \supset \bar{S}$. Define $f_S : Q \rightarrow \mathbb{R}$ by

$$f_S(x) = \begin{cases} f(x) & \text{if } x \in S, \\ 0 & \text{if } x \notin S. \end{cases} \quad (3.116)$$

By definition, f is integrable over S if and only if f_S is integrable over Q . If so then $\int_S f = \int_Q f_S$.

Let D be the set of points in Q where f_S is discontinuous. Then f_S is integrable over Q if and only if D is of measure zero. What is D ?

1. If $x \in \text{Int } S$, then $f_S = f$ in a neighborhood of x , so f_S is continuous at x .
2. If $x \in \text{Ext } S$, then $f_S = 0$ in a neighborhood of x , so f_S is continuous at x .

So, we know that $D \subseteq \text{Bd } S$. Because S is rectifiable, the boundary of S has measure zero, so D has measure zero. Thus, f_S is R. integrable, and therefore so is f . \square

Theorem 3.20. *Let A be an open subset of \mathbb{R}^n . There exists a sequence of compact rectifiable sets $C_N, N = 1, 2, 3, \dots$ such that*

$$C_N \subseteq \text{Int } C_{N+1} \quad (3.117)$$

and

$$\bigcup C_N = A. \quad (3.118)$$

Definition 3.21. The set $\{C_N\}$ is called an *exhaustion* of A .

Proof. Take the complement of A , namely $B = \mathbb{R}^n - A$. Define $d(x, B) = \inf_{y \in B} \{|x - y|\}$. The function $d(x, B)$ is a continuous function of x (the theorem for this is in section 4 of Munkres). Let

$$D_N = \{x \in A : d(x, B) \geq 1/N \text{ and } |x| \leq N\}. \quad (3.119)$$

The set D_N is compact. It is easy to check that $D_N \subseteq \text{Int } D_{N+1}$.

Claim.

$$\bigcup D_N = A. \quad (3.120)$$

Proof. Let $x \in A$. The set A is open, so there exists $\epsilon > 0$ such that the set $\{y \in \mathbb{R}^n : |y - x| \leq \epsilon\}$ is contained in A . So, $d(x, B) \geq \epsilon$.

Now, choose N such that $1/N < \epsilon$ and such that $|x| < N$. Then, by definition, $x \in D_N$. Therefore $\bigcup D_N = A$. \square

So, the D_N 's satisfy the right properties, except they are not necessarily rectifiable. We can make them rectifiable as follows.

For every $p \in D_N$, let Q_p be a rectangle with $p \in \text{Int } Q_p$ and $Q_p \subseteq \text{Int } D_{N+1}$. Then the collection of sets $\{\text{Int } Q_p : p \in D_N\}$ is an open cover of D_N . By the H-B Theorem, there exists a finite subcover $\text{Int } Q_{p_1}, \dots, \text{Int } Q_{p_r}$. Now, let

$$C_N = Q_{p_1} \cup \dots \cup Q_{p_r}. \quad (3.121)$$

Then $C_N \subseteq \text{Int } D_N \subseteq \text{Int } C_{N+1}$. \square

Lecture 13

Let A be an open set in \mathbb{R}^n , and let $f : A \rightarrow \mathbb{R}$ 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. Consider the set of all such integrals:

$$\# = \left\{ \int_D f : D \subseteq A, D \text{ compact and rectifiable} \right\}. \quad (3.122)$$

Definition 3.22. The *improper integral of f over A* exists if $\#$ is bounded, and we define the improper integral of f over A to be its l.u.b.

$$\int_A^\# f \equiv \text{l.u.b.} \int_D f = \text{improper integral of } f \text{ over } A. \quad (3.123)$$

Claim. If A is rectifiable and $f : A \rightarrow \mathbb{R}$ is bounded, then

$$\int_A^\# f = \int_A f. \quad (3.124)$$

Proof. Let $D \subseteq A$ be a compact and rectifiable set. So,

$$\int_D f \leq \int_A f \quad (3.125)$$

$$\implies \sup_D \int_D f \leq \int_A f \quad (3.126)$$

$$\implies \int_A^\# f \leq \int_A f. \quad (3.127)$$

The proof of the inequality in the other direction is a bit more complicated.

Choose a rectangle Q such that $\bar{A} \subseteq \text{Int } Q$. Define $f_A : Q \rightarrow \mathbb{R}$ by

$$f_A(x) = \begin{cases} f(x) & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \quad (3.128)$$

By definition,

$$\int_A f = \int_Q f_A. \quad (3.129)$$

Now, let P be a partition of Q , and let R_1, \dots, R_k be rectangles belonging to a partition of A . If R is a rectangle belonging to P not contained in A , then $R \cap A = \emptyset$. In such a case, $m_R(f_A) = 0$. So

$$L(f_A, P) = \sum_{i=1}^k m_{R_i}(f_A) v(R_i). \quad (3.130)$$

On the rectangle R_i ,

$$f_A = f \geq m_{R_i}(f_A). \quad (3.131)$$

So,

$$\begin{aligned} \sum_{i=1}^k m_{R_i}(f_A)v(R_i) &\leq \sum \int_{R_i} f \\ &= \int_D f \\ &\leq \int_A^\# f, \end{aligned} \quad (3.132)$$

where $D = \bigcup R_i$, which is compact and rectifiable.

The above was true for all partitions, so

$$\int_Q f_A \leq \int_Z^\# f. \quad (3.133)$$

We proved the inequality in the other direction, so

$$\int_A f = \int_A^\# f. \quad (3.134)$$

□

3.8 Exhaustions

Definition 3.23. A sequence of compact sets $C_i, i = 1, 2, 3, \dots$ is an *exhaustion of A* if $C_i \subseteq \text{Int } C_{i+1}$ for every i , and $\bigcup C_i = A$.

It is easy to see that

$$\bigcup \text{Int } C_i = A. \quad (3.135)$$

Let $C_i, i = 1, 2, 3, \dots$ be an exhaustion of A by compact rectifiable sets. Let $f : A \rightarrow \mathbb{R}$ be continuous and assume that $f \geq 0$. Note that

$$\int_{C_i} f \leq \int_{C_{i+1}} f, \quad (3.136)$$

since $C_{i+1} \supset C_i$. So

$$\int_{C_i} f, i = 1, 2, 3, \dots \quad (3.137)$$

is an increasing (actually, non-decreasing) sequence. Hence, either $\int_{C_i} f \rightarrow \infty$ as $i \rightarrow \infty$, or it has a finite limit (by which we mean $\lim_{i \rightarrow \infty} \int_{C_i} f$ exists).

Theorem 3.24. *The following two properties are equivalent:*

1. $\int_A^\# f$ exists,
2. $\lim_{i \rightarrow \infty} \int_{C_i} f$ exists.

Moreover, if either (and hence both) property holds, then

$$\int_A^\# f = \lim_{i \rightarrow \infty} \int_{C_i} f. \quad (3.138)$$

Proof. The set C_i is a compact and rectifiable set contained in A . So, if

$$\int_A^\# f \text{ exists, then} \quad (3.139)$$

$$\int_{C_i} f \leq \int_A^\# f. \quad (3.140)$$

That shows that the sets

$$\int_{C_i} f, \quad i = 1, 2, 3, \dots \quad (3.141)$$

are bounded, and

$$\lim_{i \rightarrow \infty} \int_{C_i} f \leq \int_A^\# f. \quad (3.142)$$

Now, let us prove the inequality in the other direction.

The collection of sets $\{\text{Int } C_i : i = 1, 2, 3, \dots\}$ is an open cover of A . Let $D \subseteq A$ be a compact rectifiable set contained in A . By the H-B Theorem,

$$D \subseteq \bigcup_{i=1}^N \text{Int } C_i, \quad (3.143)$$

for some N . So, $D \subseteq \text{Int } C_N \subseteq C_N$. For all such D ,

$$\int_D f \leq \int_{C_i} f \leq \lim_{i \rightarrow \infty} \int_{C_i} f. \quad (3.144)$$

Taking the infimum over all D , we get

$$\int_A^\# f \leq \lim_{i \rightarrow \infty} \int_{C_i} f. \quad (3.145)$$

We have proved the inequality in both directions, so

$$\int_A^\# f = \lim_{i \rightarrow \infty} \int_{C_i} f. \quad (3.146)$$

□

A typical illustration of this theorem is the following example.

Consider the integral

$$\int_0^1 \frac{dx}{\sqrt{x}}, \quad (3.147)$$

which we wrote in the normal integral notation from elementary calculus. In *our* notation, we would write this as

$$\int_{(0,1)} \frac{1}{\sqrt{x}}. \quad (3.148)$$

Let $C_N = [\frac{1}{N}, 1 - \frac{1}{N}]$. Then

$$\begin{aligned} \int_{(0,1)}^{\#} \frac{1}{\sqrt{x}} &= \lim_{N \rightarrow \infty} \int_{C_N} \frac{1}{\sqrt{x}} \\ &= 2\sqrt{x} \Big|_{1/N}^{1-1/N} \rightarrow 2 \text{ as } N \rightarrow \infty. \end{aligned} \quad (3.149)$$

So,

$$\int_{(0,1)}^{\#} \frac{1}{\sqrt{x}} = 2. \quad (3.150)$$

Let us now remove the assumption that $f \geq 0$. Let $f : A \rightarrow \mathbb{R}$ be any continuous function on A . As before, we define

$$f_+(x) = \max\{f(x), 0\}, \quad (3.151)$$

$$f_-(x) = \max\{-f(x), 0\}. \quad (3.152)$$

We can see that f_+ and f_- are continuous.

Definition 3.25. The improper R. integral of f over A exists if and only if the improper R. integral of f_+ and f_- over A exist. Moreover, we define

$$\int_A^{\#} f = \int_A^{\#} f_+ - \int_A^{\#} f_-. \quad (3.153)$$

We compute the integral using an exhaustion of A .

$$\begin{aligned} \int_A^{\#} f &= \lim_{N \rightarrow \infty} \left(\int_{C_N} f_+ - \int_{C_N} f_- \right) \\ &= \lim_{N \rightarrow \infty} \int_{C_N} f. \end{aligned} \quad (3.154)$$

Note that $|f| = f_+ + f_-$, so

$$\lim_{N \rightarrow \infty} \left(\int_{C_N} f_+ + \int_{C_N} f_- \right) = \lim_{N \rightarrow \infty} \int_{C_N} |f|. \quad (3.155)$$

Therefore, the improper integral of f exists if and only if the improper integral of $|f|$ exists.

Define a function $f : \mathbb{R} \rightarrow \mathbb{R}$ by

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ e^{-1/x} & \text{if } x > 0. \end{cases} \quad (3.156)$$

This is a $\mathcal{C}^\infty(\mathbb{R})$ function. Clearly, $f'(x) = f''(x) = \dots = 0$ when $x = 0$, so in the Taylor series expansion of f at zero,

$$\sum a_n x^n = 0, \quad (3.157)$$

all of the coefficients a_n are zero. However, f has a non-zero value in every neighborhood of zero.

Take $a \in \mathbb{R}$ and $\epsilon > 0$. Define a new function $g_{a,a+\epsilon} : \mathbb{R} \rightarrow \mathbb{R}$ by

$$g_{a,a+\epsilon}(x) = \frac{f(x-a)}{f(x-a) + f(a+\epsilon-x)}. \quad (3.158)$$

The function $g_{a,a+\epsilon}$ is a $\mathcal{C}^\infty(\mathbb{R})$ function. Notice that

$$g_{a,a+\epsilon} = \begin{cases} 0 & \text{if } x \leq a, \\ 1 & \text{if } x \geq a + \epsilon. \end{cases} \quad (3.159)$$

Take b such that $a < a + \epsilon < b - \epsilon < b$. Define a new function $h_{a,b} \in \mathcal{C}^\infty(\mathbb{R})$ by

$$h_{a,b}(x) = g_{a,a+\epsilon}(x)(1 - g_{a-\epsilon,b}(x)). \quad (3.160)$$

Notice that

$$h_{a,b} = \begin{cases} 0 & \text{if } x \leq a, \\ 1 & \text{if } a + \epsilon \leq x \leq b - \epsilon, \\ 0 & \text{if } b \leq x. \end{cases} \quad (3.161)$$

Lecture 14

As before, let $f : \mathbb{R} \rightarrow \mathbb{R}$ be the map defined by

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ e^{-1/x} & \text{if } x > 0. \end{cases} \quad (3.162)$$

This is a *Cinf*(\mathbb{R}) function. Take the interval $[a, b] \in \mathbb{R}$ and define the function $f_{a,b} : \mathbb{R} \rightarrow \mathbb{R}$ by $f_{a,b}(x) = f(x-a)f(b-x)$. Note that $f_{a,b} > 0$ on (a, b) , and $f_{a,b} = 0$ on $\mathbb{R} - (a, b)$.

We generalize the definition of f to higher dimensions. Let $Q \subseteq \mathbb{R}^n$ be a rectangle, where $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$. Define a new map $f_Q : \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$f_Q(x_1, \dots, x_n) = f_{a_1, b_1}(x_1) \cdots f_{a_n, b_n}(x_n). \quad (3.163)$$

Note that $f_Q > 0$ on $\text{Int } Q$, and that $f_Q = 0$ on $\mathbb{R}^n - \text{Int } Q$.

3.9 Support and Compact Support

Now for some terminology. Let U be an open set in \mathbb{R}^n , and let $f : U \rightarrow \mathbb{R}$ be a continuous function.

Definition 3.26. The *support* of f is

$$\text{supp } f = \overline{\{x \in U : f(x) \neq 0\}}. \quad (3.164)$$

For example, $\text{supp } f_Q = Q$.

Definition 3.27. Let $f : U \rightarrow \mathbb{R}$ be a continuous function. The function f is *compactly supported* if $\text{supp } f$ is compact.

Notation.

$$\mathcal{C}_0^k(U) = \text{The set of compactly supported } \mathcal{C}^k \text{ functions on } U. \quad (3.165)$$

Suppose that $f \in \mathcal{C}_0^k(U)$. Define a new set $U_1 = (\mathbb{R}^n - \text{supp } f)$. Then $U \cup U_1 = \mathbb{R}^n$, because $\text{supp } f \subseteq U$.

Define a new map $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$\tilde{f} = \begin{cases} f & \text{on } U, \\ 0 & \text{on } U_1. \end{cases} \quad (3.166)$$

The function \tilde{f} is \mathcal{C}^k on U and \mathcal{C}^k on U_1 , so \tilde{f} is in $\mathcal{C}_0^k(\mathbb{R}^n)$.

So, whenever we have a function $f \in \mathcal{C}^k$ is compactly supported on U , we can drop the tilde and think of f as in $\mathcal{C}_0^k(\mathbb{R}^n)$.

3.10 Partitions of Unity

Let $\{U_\alpha : \alpha \in I\}$ be a collection of open subsets of \mathbb{R}^n such that $U = \cup_\alpha U_\alpha$.

Theorem 3.28. *There exists a sequence of rectangles Q_i , $i = 1, 2, 3, \dots$ such that*

1. $\text{Int } Q_i$, $i = 1, 2, 3, \dots$ is a cover of U ,
2. Each $Q_i \subset U_\alpha$ for some α ,
3. For every point $p \in U$, there exists a neighborhood U_p of p such that $U_p \cap Q_i = \emptyset$ for all $i > N_p$.

Proof. Take an exhaustion A_1, A_2, A_3, \dots of U . By definition, the exhaustion satisfies

$$\begin{cases} A_i \subseteq \text{Int } A_{i+1} \\ A_i \text{ is compact} \\ \cup A_i = U. \end{cases}$$

We previously showed that you can always find an exhaustion.

Let $B_i = A_i - \text{Int } A_{i-1}$. For each $x \in B_i$, let Q_x be a rectangle with $x \in \text{Int } Q_x$ such that $Q_x \subseteq U_\alpha$, for some alpha, and $Q_x \subset \text{Int } A_{i+1} - A_{i-2}$. Then, the collection of sets $\{\text{Int } Q_x : x \in B_i\}$ covers B_i . Each set B_i is compact, so, by the H-B Theorem, there exists a finite subcover $\text{Int } Q_{x_r} \equiv \text{Int } Q_{i,r}$, $r = 1, \dots, N_i$.

The rectangles $Q_{i,r}$, $1 \leq r \leq N_i$, $i = 1, 2, 3, \dots$ satisfy the hypotheses of the theorem, after relabeling the rectangles in linear sequence Q_1, Q_2, Q_3 , etc. (you should check this). \square

The following theorem is called the Partition of Unity Theorem.

Theorem 3.29. *There exist functions $f_i \in C_0^\infty(U)$ such that*

1. $f_i \geq 0$,
2. $\text{supp } f_i \subseteq U_\alpha$, for some α ,
3. For every $p \in U$, there exists a neighborhood U_p of p such that $U_p \cap \text{supp } f_i = \emptyset$ for all $i > N_p$,
4. $\sum f_i = 1$.

Proof. Let Q_i , $i = 1, 2, 3, \dots$ be a collection of rectangles with the properties of the previous theorem. Then the functions f_{Q_i} , $i = 1, 2, 3, \dots$ have all the properties presented in the theorem, except for property 4. We now prove the fourth property. We now that $f_{Q_i} > 0$ on $\text{Int } Q_i$, and $\{\text{Int } Q_i : i = 1, 2, 3, \dots\}$ is a cover of U . So, for every $p \in U$, $f_{Q_i}(p) > 0$ for some i . So

$$\sum f_{Q_i} > 0. \tag{3.167}$$

We can divide by a nonzero number, so we can define

$$f_i = \frac{f_{Q_i}}{\sum_{i=1}^{\infty} f_{Q_i}}. \quad (3.168)$$

This new function satisfies property 4. Note that the infinite sum converges because the sum has only a finite number of nonzero terms. \square

Lecture 15

We restate the partition of unity theorem from last time. Let $\{U_\alpha : \alpha \in I\}$ be a collection of open subsets of \mathbb{R}^n such that

$$U = \bigcup_{\alpha \in I} U_\alpha. \quad (3.169)$$

Theorem 3.30. *There exist functions $f_i \subseteq C_0^\infty(U)$ such that*

1. $f_i \geq 0$,
2. $\text{supp } f_i \subseteq U_\alpha$, for some α ,
3. For every $p \in U$, there exists a neighborhood U_p of p such that $U_p \cap \text{supp } f_i = \emptyset$ for all $i > N_p$,
4. $\sum f_i = 1$.

Remark. Property (4) makes sense because of property (3), because at each point it is a finite sum.

Remark. A set of functions satisfying properties (1), (3), and (4) is called a *partition of unity*.

Remark. Property (2) can be restated as “the partition of unity is subordinate to the cover $\{U_\alpha : \alpha \in I\}$.”

Let us look at some typical applications of partitions of unity.

The first application is to improper integrals. Let $\phi : U \rightarrow \mathbb{R}$ be a continuous map, and suppose

$$\int_U \phi \quad (3.170)$$

is well-defined. Take a partition of unity $\sum f_i = 1$. The function $f_i \phi$ is continuous and compactly supported, so it is bounded. Let $\text{supp } f_i \subseteq Q_i$ for some rectangle Q_i . Then,

$$\int_{Q_i} f_i \phi \quad (3.171)$$

is a well-defined R. integral. It follows that

$$\int_U f_i \phi = \int_{Q_i} f_i \phi. \quad (3.172)$$

It follows that

$$\int_U \phi = \sum_{i=1}^{\infty} \int_{Q_i} f_i \phi. \quad (3.173)$$

This is proved in Munkres.

The second application of partitions of unity involves *cut-off functions*.

Let $f_i \in C_0^\infty(U)$, $i = 1, 2, 3, \dots$ be a partition of unity, and let $A \subseteq U$ be compact.

Lemma 3.31. *There exists a neighborhood U' of A in U and a number $N > 0$ such that $A \cup \text{supp } f_i = \phi$ for all $i > N$.*

Proof. For any $p \in A$, there exists a neighborhood U_p of p and a number N_p such that $U' \cup \text{supp } f_i = \phi$ for all $i > N_p$. The collection of all these U_p is a cover of A . By the H-B Theorem, there exists a finite subcover U_{p_i} , $i = 1, 2, 3, \dots$ of A . Take $U_p = \cup U_{p_i}$ and take $N = \max\{N_{p_i}\}$. \square

We use this lemma to prove the following theorem.

Theorem 3.32. *Let $A \subseteq \mathbb{R}^n$ be compact, and let U be an open set containing A . There exists a function $f \in C_0^\infty(U)$ such that $f \equiv 1$ (identically equal to 1) on a neighborhood $U' \subset U$ of A .*

Proof. Choose U' and N as in the lemma, and let

$$f = \sum_{i=1}^N f_i. \quad (3.174)$$

Then $\text{supp } f_i \cap U' = \phi$ for all $i > N$. So, on U' ,

$$f = \sum_{i=1}^{\infty} f_i = 1. \quad (3.175)$$

\square

Such an f can be used to create cut-off functions. We look at an application.

Let $\phi : U \rightarrow \mathbb{R}$ be a continuous function. Define $\psi = f\phi$. The new function ψ is called a cut-off function, and it is compactly supported with $\text{supp } \phi \subseteq U$. We can extend the domain of ψ by defining $\psi = 0$ outside of U . The extended function $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ is still continuous, and it equals ϕ on a neighborhood of A .

We look at another application, this time to *exhaustion functions*.

Definition 3.33. Given an open set U , and a collection of compact subsets A_i $i = 1, 2, 3, \dots$ of U , the sets A_i form an *exhaustion of U* if $A_i \subseteq \text{Int } A_{i+1}$ and $\cup A_i = U$ (this is just a quick reminder of the definition of exhaustion).

Definition 3.34. A function $\phi \in C^\infty(U)$ is an *exhaustion function* if

1. $\phi > 0$,
2. the sets $A_i = \phi^{-1}([0, 1])$ are compact.

Note that this implies that the A_i 's are an exhaustion.

We use the fact that we can always find a partition of unity to show that we can always find exhaustion functions.

Take a partition of unity $f_i \in C^\infty(U)$, and define

$$\phi = \sum_{i=1}^{\infty} i f_i. \quad (3.176)$$

This sum converges because only finitely many terms are nonzero.

Consider any point

$$p \notin \bigcup_{j \leq i} \text{supp } f_j. \quad (3.177)$$

Then,

$$\begin{aligned} 1 &= \sum_{k=1}^{\infty} f_k(p) \\ &= \sum_{k>i} f_k(p), \end{aligned} \quad (3.178)$$

so

$$\begin{aligned} \sum_{\ell=1}^{\infty} \ell f_\ell(p) &= \sum_{\ell>i} \ell f_\ell \\ &\geq i \sum_{\ell>i} f_\ell \\ &= i. \end{aligned} \quad (3.179)$$

That is, if $p \notin \bigcup_{j \leq i} \text{supp } f_j$, then $f(p) > i$. So,

$$\phi^{-1}([0, i]) \subseteq \bigcup_{j \leq i} \text{supp } f_j, \quad (3.180)$$

which you should check yourself. The compactness of the r.h.s. implies the compactness of the l.h.s.

Now we look at problem number 4 in section 16 of Munkres. Let A be an arbitrary subset of \mathbb{R}^n , and let $g : A \rightarrow \mathbb{R}^k$ be a map.

Definition 3.35. The function g is C^k on A if for every $p \in A$, there exists a neighborhood U_p of p in \mathbb{R}^n and a C^k map $g^p : U_p \rightarrow \mathbb{R}^k$ such that $g^p|_{U_p \cap A} = g|_{U_p \cap A}$.

Theorem 3.36. If $g : A \rightarrow \mathbb{R}^k$ is C^k , then there exists a neighborhood U of A in \mathbb{R}^n and a C^k map $\tilde{g} : U \rightarrow \mathbb{R}^k$ such that $\tilde{g} = g$ on A .

Proof. This is a very nice application of partition of unity. Read Munkres for the proof. \square

Lecture 16

4 Multi-linear Algebra

4.1 Review of Linear Algebra and Topology

In today's lecture we review chapters 1 and 2 of Munkres. Our ultimate goal (not today) is to develop vector calculus in n dimensions (for example, the generalizations of grad, div, and curl).

Let V be a vector space, and let $v_i \in V, i = 1, \dots, k$.

1. The v_i 's are *linearly independent* if the map from \mathbb{R}^k to V mapping (c_1, \dots, c_k) to $c_1v_1 + \dots + c_kv_k$ is injective.
2. The v_i 's *span* V if this map is surjective (onto).
3. If the v_i 's form a basis, then $\dim V = k$.
4. A subset W of V is a *subspace* if it is also a vector space.
5. Let V and W be vector spaces. A map $A : V \rightarrow W$ is *linear* if $A(c_1v_1 + c_2v_2) = c_1A(v_1) + c_2A(v_2)$.
6. The *kernel* of a linear map $A : V \rightarrow W$ is

$$\ker A = \{v \in V : Av = 0\}. \quad (4.1)$$

7. The *image* of A is

$$\text{Im } A = \{Av : v \in V\}. \quad (4.2)$$

8. The following is a basic identity:

$$\dim \ker A + \dim \text{Im } A = \dim V. \quad (4.3)$$

9. We can associate linear mappings with matrices. Let v_1, \dots, v_n be a basis for V , and let w_1, \dots, w_m be a basis for W . Let

$$Av_j = \sum_{i=1}^m a_{ij}w_i. \quad (4.4)$$

Then we associate the linear map A with the matrix $[a_{ij}]$. We write this $A \sim [a_{ij}]$.

10. If v_1, \dots, v_n is a basis for V and $u_j = \sum a_{ij}w_i$ are n arbitrary vectors in W , then there exists a unique linear mapping $A : V \rightarrow W$ such that $Av_j = u_j$.

11. Know all the material in Munkres section § 2 on matrices and determinants.
12. The quotient space construction. Let V be a vector space and W a subspace. Take any $v \in V$. We define $v + W \equiv \{v + w : w \in W\}$. Sets of this form are called W -cosets. One can check that given $v_1 + W$ and $v_2 + W$,
- (a) If $v_1 - v_2 \in W$, then $v_1 + W = v_2 + W$.
 - (b) If $v_1 - v_2 \notin W$, then $(v_1 + W) \cap (v_2 + W) = \emptyset$.

So every vector $v \in V$ belongs to a unique W -coset.

The *quotient space* V/W is the set of all W -cosets.

For example, let $V = \mathbb{R}^2$, and let $W = \{(a, 0) : a \in \mathbb{R}\}$. The W -cosets are then vertical lines.

The set V/W is a vector space. It satisfies vector addition: $(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W$. It also satisfies scalar multiplication: $\lambda(v + W) = \lambda v + W$. You should check that the standard axioms for vector spaces are satisfied.

There is a natural projection from V to V/W :

$$\pi : V \rightarrow V/W, v \rightarrow v + W. \quad (4.5)$$

The map π is a linear map, it is surjective, and $\ker \pi = W$. Also, $\text{Im } \pi = V/W$, so

$$\begin{aligned} \dim V/W &= \dim \text{Im } \pi \\ &= \dim V - \dim \ker \pi \\ &= \dim V - \dim W. \end{aligned} \quad (4.6)$$

4.2 Dual Space

13. The *dual space* construction: Let V be an n -dimensional vector space. Define V^* to be the set of all linear functions $\ell : V \rightarrow \mathbb{R}$. Note that if $\ell_1, \ell_2 \in V^*$ and $\lambda_1, \lambda_2 \in \mathbb{R}$, then $\lambda_1 \ell_1 + \lambda_2 \ell_2 \in V^*$, so V^* is a vector space.

What does V^* look like? Let e_1, \dots, e_n be a basis of V . By item (9), there exists a unique linear map $e_i^* \in V^*$ such that

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

Claim. *The set of vectors e_1^*, \dots, e_n^* is a basis of V^* .*

Proof. Suppose $\ell = \sum c_i e_i^* = 0$. Then $0 = \ell(e_j) = \sum c_i e_i^*(e_j) = c_j$, so $c_1 = \dots = c_n = 0$. This proves that the vectors e_i^* are linearly independent. Now, if $\ell \in V^*$ and $\ell(e_i) = c_j$ one can check that $\ell = \sum c_i e_i^*$. This proves that the vectors e_i^* span V^* . \square

The vectors e_1^*, \dots, e_n^* are said to be a *basis of V^* dual to e_1, \dots, e_n* .

Note that $\dim V^* = \dim V$.

Suppose that we have a pair of vectors spaces V, W and a linear map $A : V \rightarrow W$. We get another map

$$A^* : W^* \rightarrow V^*, \quad (4.7)$$

defined by $A^*\ell = \ell \circ A$, where $\ell \in W^*$ is a linear map $\ell : W \rightarrow \mathbb{R}$. So $A^*\ell$ is a linear map $A^*\ell : V \rightarrow \mathbb{R}$. You can check that $A^* : W^* \rightarrow V^*$ is linear.

We look at the matrix description of A^* . Define the following bases:

$$e_1, \dots, e_n \text{ a basis of } V \quad (4.8)$$

$$f_1, \dots, f_n \text{ a basis of } W \quad (4.9)$$

$$e_1^*, \dots, e_n^* \text{ a basis of } V^* \quad (4.10)$$

$$f_1^*, \dots, f_n^* \text{ a basis of } W^*. \quad (4.11)$$

Then

$$\begin{aligned} A^*f_j^*(e_i) &= f_j^*(Ae_i) \\ &= f_j^*\left(\sum_k a_{ki}f_k\right) \\ &= a_{ji} \end{aligned} \quad (4.12)$$

So,

$$A^*f_j = \sum_k a_{jk}e_k^*, \quad (4.13)$$

which shows that $A^* \sim [a_{ji}] = [a_{ij}]^t$, the transpose of A .

Lecture 17

Today we begin studying the material that is also found in the Multi-linear Algebra Notes. We begin with the theory of *tensors*.

4.3 Tensors

Let V be a n -dimensional vector space. We use the following notation.

Notation.

$$V^k = \underbrace{V \times \cdots \times V}_{k \text{ times}}. \quad (4.14)$$

For example,

$$V^2 = V \times V, \quad (4.15)$$

$$V^3 = V \times V \times V. \quad (4.16)$$

Let $T : V^k \rightarrow \mathbb{R}$ be a map.

Definition 4.1. The map T is *linear in its i th factor* if for every sequence $v_j \in V, 1 \leq j \leq n, j \neq i$, the function mapping $v \in V$ to $T(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_k)$ is linear in v .

Definition 4.2. The map T is *k -linear* (or is a *k -tensor*) if it is linear in all k factors.

Let T_1, T_2 be k -tensors, and let $\lambda_1, \lambda_2 \in \mathbb{R}$. Then $\lambda_1 T_1 + \lambda_2 T_2$ is a k -tensor (it is linear in all of its factors).

So, the set of all k -tensors is a vector space, denoted by $\mathcal{L}^k(V)$, which we sometimes simply denote by \mathcal{L}^k .

Consider the special case $k = 1$. The the set $\mathcal{L}^1(V)$ is the set of all linear maps $\ell : V \rightarrow \mathbb{R}$. In other words,

$$\mathcal{L}^1(V) = V^*. \quad (4.17)$$

We use the convention that

$$\mathcal{L}^0(V) = \mathbb{R}. \quad (4.18)$$

Definition 4.3. Let $T_i \in \mathcal{L}^{k_i}, i = 1, 2$, and define $k = k_1 + k_2$. We define the *tensor product* of T_1 and T_2 to be the tensor $T_1 \otimes T_2 : V^k \rightarrow \mathbb{R}$ defined by

$$T_1 \otimes T_2(v_1, \dots, v_k) = T_1(v_1, \dots, v_{k_1})T_2(v_{k_1+1}, \dots, v_k). \quad (4.19)$$

We can conclude that $T_1 \otimes T_2 \in \mathcal{L}^k$.

We can define more complicated tensor products. For example, let $T_i \in \mathcal{L}^{k_i}, i = 1, 2, 3$, and define $k = k_1 + k_2 + k_3$. Then we have the tensor product

$$\begin{aligned} T_1 \otimes T_2 \otimes T_3(v_1, \dots, v_k) \\ = T_1(v_1, \dots, v_{k_1})T_2(v_{k_1+1}, \dots, v_{k_1+k_2})T_3(v_{k_1+k_2+1}, \dots, v_k). \end{aligned} \quad (4.20)$$

Then $T_1 \otimes T_2 \otimes T_3 \in \mathcal{L}^k$. Note that we could have simply defined

$$\begin{aligned} T_1 \otimes T_2 \otimes T_3 &= (T_1 \otimes T_2) \otimes T_3 \\ &= T_1 \otimes (T_2 \otimes T_3), \end{aligned} \tag{4.21}$$

where the second equality is the associative law for tensors. There are other laws, which we list here.

- Associative Law: $(T_1 \otimes T_2) \otimes T_3 = T_1 \otimes (T_2 \otimes T_3)$.
- Right and Left Distributive Laws: Suppose $T_i \in \mathcal{L}^{k_i}, i = 1, 2, 3$, and assume that $k_1 = k_2$. Then
 - Left: $(T_1 + T_2) \otimes T_3 = T_1 \otimes T_3 + T_2 \otimes T_3$.
 - Right: $T_3 \otimes (T_1 + T_2) = T_3 \otimes T_1 + T_3 \otimes T_2$.
- Let λ be a scalar. Then

$$\lambda(T_1 \otimes T_2) = (\lambda T_1) \otimes T_2 = T_1 \otimes (\lambda T_2). \tag{4.22}$$

Now we look at an important class of k -tensors. Remember that $\mathcal{L}^1(V) = V^*$, and take any 1-tensors $\ell_i \in V^*, i = 1, \dots, k$.

Definition 4.4. The tensor $T = \ell_1 \otimes \dots \otimes \ell_k$ is a *decomposable k -tensor*.

By definition, $T(v_1, \dots, v_k) = \ell_1(v_1) \dots \ell_k(v_k)$. That is, $\ell_1 \otimes \dots \otimes \ell_k(v_1, \dots, v_k) = \ell_1(v_1) \dots \ell_k(v_k)$.

Now let us go back to considering $\mathcal{L}^k = \mathcal{L}^k(V)$.

Theorem 4.5.

$$\dim \mathcal{L}^k = n^k. \tag{4.23}$$

Note that for $k = 1$, this shows that $\mathcal{L}^1(V) = V^$ has dimension n .*

Proof. Fix a basis e_1, \dots, e_n of V . This defines a dual basis e_1^*, \dots, e_n^* of V^* , $e_i^* : V \rightarrow \mathbb{R}$ defined by

$$e_i^*(e_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{4.24}$$

Definition 4.6. A *multi-index* I of length k is a set of integers $(i_1, \dots, i_k), 1 \leq i_r \leq n$. We define

$$e_I^* = e_{i_1}^* \otimes \dots \otimes e_{i_k}^* \in \mathcal{L}^k. \tag{4.25}$$

Let $J = (j_1, \dots, j_k)$ be a multi-index of length k . Then

$$e_I^*(e_{j_1}, \dots, e_{j_k}) = e_{i_1}^*(e_{j_1}) \dots e_{i_k}^*(e_{j_k}) = \begin{cases} 1 & \text{if } I = J, \\ 0 & \text{if } I \neq J. \end{cases} \tag{4.26}$$

Claim. The k -tensors e_I^* are a basis of \mathcal{L}^k .

Proof. To prove the claim, we use the following lemma.

Lemma 4.7. Let T be a k -tensor. Suppose that $T(e_{i_1}, \dots, e_{i_k}) = 0$ for all multi-indices I . Then $T = 0$.

Proof. Define a $(k-1)$ -tensor $T_i : V^{k-1} \rightarrow \mathbb{R}$ by setting

$$T_i(v_1, \dots, v_{k-1}) = T(v_1, \dots, v_{k-1}, e_i), \quad (4.27)$$

and let $v_k = \sum a_i e_i$. By linearity, $T(v_1, \dots, v_k) = \sum a_i T_i(v_1, \dots, v_{k-1})$. So, if the lemma is true for the T_i 's, then it is true for T by an induction argument (we leave this to the student to prove). \square

With this lemma we can prove the claim.

First we show that the e_I^* 's are linearly independent. Suppose that

$$0 = T = \sum c_I e_I^*. \quad (4.28)$$

For any multi-index J of length k ,

$$\begin{aligned} 0 &= T(e_{j_1}, \dots, e_{j_k}) \\ &= \sum c_I e_I^*(e_{j_1}, \dots, e_{j_k}) \\ &= c_J \\ &= 0. \end{aligned} \quad (4.29)$$

So the e_I^* 's are linearly independent.

Now we show that the e_I^* 's span \mathcal{L}^k . Let $T \in \mathcal{L}^k$. For every I let $T_I = T(e_{i_1}, \dots, e_{i_k})$, and let $T' = \sum T_I e_I^*$. One can check that $(T - T')(e_{j_1}, \dots, e_{j_k}) = 0$ for all multi-indices J . Then the lemma tells us that $T = T'$, so the e_I^* 's span \mathcal{L}^k , which proves our claim. \square

Since the e_I^* 's are a basis of \mathcal{L}^k , we see that

$$\dim \mathcal{L}^k = n^k, \quad (4.30)$$

which proves our theorem. \square

4.4 Pullback Operators

Let V, W be vector spaces, and let $A : V \rightarrow W$ be a linear map. Let $T \in \mathcal{L}^k(W)$, and define a new map $A^*T \in \mathcal{L}^k(V)$ (called the “pullback” tensor) by

$$A^*T(v_1, \dots, v_k) = T(Av_1, \dots, Av_k). \quad (4.31)$$

You should prove the following claims as an exercise:

Claim. *The map $A^* : \mathcal{L}^k(W) \rightarrow \mathcal{L}^k(V)$ is a linear map.*

Claim. *Let $T_i \in \mathcal{L}^{k_i}(W)$, $i = 1, 2$. Then*

$$A^*(T_1 \otimes T_2) = A^*T_1 \otimes A^*T_2. \quad (4.32)$$

Now, let $A : V \rightarrow W$ and $B : W \rightarrow U$ be maps, where U is a vector space. Given $T \in \mathcal{L}^k(U)$, we can “pullback” to W by B^*T , and then we can “pullback” to V by $A^*(B^*T) = (B \circ A)^*T$.

4.5 Alternating Tensors

In this course we will be restricting ourselves to *alternating tensors*.

Definition 4.8. A permutation of order k is a bijective map

$$\sigma : \{1, \dots, k\} \rightarrow \{1, \dots, k\}. \quad (4.33)$$

The map is a bijection, so σ^{-1} exists.

Given two permutations σ_1, σ_2 , we can construct the composite permutation

$$\sigma_1 \circ \sigma_2(i) = \sigma_1(\sigma_2(i)). \quad (4.34)$$

We define

$$S_k \equiv \text{The set of all permutations of } \{1, \dots, k\}. \quad (4.35)$$

There are some special permutations. Fix $1 \leq i < j \leq k$. Let τ be the permutation such that

$$\tau(i) = j \quad (4.36)$$

$$\tau(j) = i \quad (4.37)$$

$$\tau(\ell) = \ell, \ell \neq i, j. \quad (4.38)$$

The permutation τ is called a *transposition*.

Definition 4.9. The permutation τ is an elementary transposition if $j = i + 1$.

We state without proof two very useful theorems.

Theorem 4.10. *Every permutation can be written as a product $\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_m$, where each τ_i is an elementary transposition.*

Theorem 4.11. *Every permutation σ can be written either as a product of an even number of elementary transpositions or as a product of an odd number of elementary transpositions, but not both.*

Because of the second theorem, we can define an important invariant of a permutation: the *sign of the permutation*.

Definition 4.12. If $\sigma = \tau_1 \circ \cdots \circ \tau_m$, where the τ_i 's are elementary transpositions, then the *sign of σ* is

$$\text{sign of } \sigma = (-1)^\sigma = (-1)^m. \quad (4.39)$$

Note that if $\sigma = \sigma_1 \circ \sigma_2$, then $(-1)^\sigma = (-1)^{\sigma_1}(-1)^{\sigma_2}$. We can see this by letting $\sigma_1 = \tau_1 \circ \cdots \circ \tau_{m_1}$, and $\sigma_2 = \tau'_1 \circ \cdots \circ \tau'_{m_2}$, and noting that $\sigma_1 \circ \sigma_2 = \tau_1 \circ \cdots \circ \tau_{m_1} \circ \tau'_1 \circ \cdots \circ \tau'_{m_2}$.

Lecture 18

We begin with a quick review of permutations (from last lecture).

A permutation of order k is a bijective map $\sigma : \{1, \dots, k\} \rightarrow \{1, \dots, k\}$. We denote by S_k the set of permutations of order k .

The set S_k has some nice properties. If $\sigma \in S_k$, then $\sigma^{-1} \in S_k$. The inverse permutation σ^{-1} is defined by $\sigma^{-1}(j) = i$ if $\sigma(i) = j$. Another nice property is that if $\sigma, \tau \in S_k$, then $\sigma\tau \in S_k$, where $\sigma\tau(i) = \sigma(\tau(i))$. That is, if $\tau(i) = j$ and $\sigma(j) = k$, then $\sigma\tau(i) = k$.

Take $1 \leq i < j \leq k$, and define

$$\tau_{i,j}(i) = j \tag{4.40}$$

$$\tau_{i,j}(j) = i \tag{4.41}$$

$$\tau_{i,j}(\ell) = \ell, \ell \neq i, j. \tag{4.42}$$

The permutation $\tau_{i,j}$ is a transposition. It is an elementary transposition of $j = i + 1$. Last time we stated the following theorem.

Theorem 4.13. *Every permutation σ can be written as a product*

$$\sigma = \tau_1 \tau_2 \cdots \tau_r, \tag{4.43}$$

where the τ_i 's are elementary transpositions.

In the above, we removed the symbol \circ denoting composition of permutations, but the composition is still implied.

Last time we also defined the sign of a permutation

Definition 4.14. The sign of a permutation σ is $(-1)^\sigma = (-1)^r$, where r is as in the above theorem.

Theorem 4.15. *The above definition of sign is well-defined, and*

$$(-1)^{\sigma\tau} = (-1)^\sigma (-1)^\tau. \tag{4.44}$$

All of the above is discussed in the Multi-linear Algebra Notes.

Part of today's homework is to show the following two statements:

1. $|S_k| = k!$. The proof is by induction.
2. $(-1)^{\tau_{i,j}} = -1$. Hint: use induction and $\tau_{i,j} = (\tau_{j-1,j})(\tau_{i,j-1})(\tau_{j-1,j})$, with $i < j$.

We now move back to the study of tensors. Let V be an n -dimensional vector space. We define

$$V^k = \underbrace{V \times \cdots \times V}_{k \text{ factors}}. \tag{4.45}$$

We define $\mathcal{L}^k(v)$ to be the space of all k -linear functions $T : V^k \rightarrow \mathbb{R}$. If $T_i \in \mathcal{L}^{k_i}$, $i = 1, 2$, and $k = k_1 + k_2$, then $T_1 \otimes T_2 \in \mathcal{L}^k$. Decomposable k -tensors are of the form $T = \ell_1 \otimes \cdots \otimes \ell_k$, where each $\ell_i \in \mathcal{L}^1 = V^*$. Note that $\ell_1 \otimes \cdots \otimes \ell_k(v_1, \dots, v_k) = \ell_1(v_1) \cdots \ell_k(v_k)$.

We define a permutation operation on tensors. Take $\sigma \in S_k$ and $T \in \mathcal{L}^k(V)$.

Definition 4.16. We define the map $T^\sigma : V^k \rightarrow \mathbb{R}$ by

$$T^\sigma(v_1, \dots, v_k) = T(v_{\sigma^{-1}(1)}, \dots, v_{\sigma^{-1}(k)}). \quad (4.46)$$

Clearly, $T^\sigma \in \mathcal{L}^k(V)$. We have the following useful formula:

Claim.

$$(T^\sigma)^\tau = T^{\tau\sigma}. \quad (4.47)$$

Proof.

$$\begin{aligned} T^{\tau\sigma}(v_1, \dots, v_k) &= T(v_{\sigma^{-1}(\tau^{-1}(1))}, \dots, v_{\sigma^{-1}(\tau^{-1}(k))}) \\ &= T^\sigma(v_{\tau^{-1}(1)}, \dots, v_{\tau^{-1}(k)}) \\ &= (T^\sigma)^\tau(v_1, \dots, v_k). \end{aligned} \quad (4.48)$$

□

Let us look at what the permutation operation does to a decomposable tensor $T = \ell_1 \otimes \cdots \otimes \ell_k$.

$$T^\sigma(v_1, \dots, v_k) = \ell_1(v_{\sigma^{-1}(1)}) \cdots \ell_k(v_{\sigma^{-1}(k)}). \quad (4.49)$$

The i th factor has the subscript $\sigma^{-1}(i) = j$, where $\sigma(j) = i$, so the i th factor is $\ell_{\sigma(j)}(v_j)$. So

$$\begin{aligned} T^\sigma(v_1, \dots, v_k) &= \ell_{\sigma(1)}(v_1) \cdots \ell_{\sigma(k)}(v_k) \\ &= (\ell_{\sigma(1)} \otimes \cdots \otimes \ell_{\sigma(k)})(v_1, \dots, v_k). \end{aligned} \quad (4.50)$$

To summarize,

$$\begin{cases} T = \ell_1 \otimes \cdots \otimes \ell_k \\ T^\sigma = \ell_{\sigma(1)} \otimes \cdots \otimes \ell_{\sigma(k)}. \end{cases} \quad (4.51)$$

Proposition 4.17. *The mapping $T \in \mathcal{L}^k \rightarrow T^\sigma \in \mathcal{L}^k$ is linear.*

We leave the proof of this as an exercise.

Definition 4.18. A tensor $T \in \mathcal{L}^k(V)$ is *alternating* if $T^\sigma = (-1)^\sigma T$ for all $\sigma \in S_k$.

Definition 4.19. We define

$$\mathcal{A}^k(V) = \text{the set of all alternating } k\text{-tensors}. \quad (4.52)$$

By our previous claim, \mathcal{A}^k is a vector space.

The *alternating operator* Alt can be used to create alternating tensors.

Definition 4.20. Given a k -tensor $T \in \mathcal{L}^k(V)$, we define the *alternating operator* $\text{Alt} : \mathcal{L}^k(V) \rightarrow \mathcal{A}^k(V)$ by

$$\text{Alt}(T) = \sum_{\tau \in S_k} (-1)^\tau T^\tau. \quad (4.53)$$

Claim. *The alternating operator has the following properties:*

1. $\text{Alt}(T) \in \mathcal{A}^k(V)$,
2. If $T \in \mathcal{A}^k(V)$, then $\text{Alt}(T) = k!T$,
3. $\text{Alt}(T^\sigma) = (-1)^\sigma \text{Alt}(T)$,
4. The map $\text{Alt} : \mathcal{L}^k(V) \rightarrow \mathcal{A}^k(V)$ is linear.

Proof. 1.

$$\text{Alt}(T) = \sum_{\tau} (-1)^\tau T^\tau, \quad (4.54)$$

so

$$\begin{aligned} \text{Alt}(T)^\sigma &= \sum_{\tau} (-1)^\tau (T^\tau)^\sigma \\ &= \sum_{\tau} (-1)^\tau T^{\sigma\tau} \\ &= (-1)^\sigma \sum_{\sigma\tau} (-1)^{\sigma\tau} T^{\sigma\tau} \\ &= (-1)^\sigma \text{Alt}(T). \end{aligned} \quad (4.55)$$

2.

$$\text{Alt}(T) = \sum_{\tau} (-1)^\tau T^\tau, \quad (4.56)$$

but $T^\tau = (-1)^\tau T$, since $T \in \mathcal{A}^k(V)$. So

$$\begin{aligned} \text{Alt}(T) &= \sum_{\tau} (-1)^\tau (-1)^\tau T \\ &= k!T. \end{aligned} \quad (4.57)$$

3.

$$\begin{aligned} \text{Alt}(T^\sigma) &= \sum_{\tau} (-1)^\tau (T^\sigma)^\tau \\ &= \sum_{\tau} (-1)^\tau T^{\tau\sigma} \\ &= (-1)^\sigma \sum_{\tau\sigma} (-1)^{\tau\sigma} T^{\tau\sigma} \\ &= (-1)^\sigma \text{Alt}(T). \end{aligned} \quad (4.58)$$

4. We leave the proof as an exercise. □

Now we ask ourselves: what is the dimension of $\mathcal{A}^k(V)$? To answer this, it is best to write a basis.

Earlier we found a basis for \mathcal{L}^k . We defined e_1, \dots, e_n to be a basis of V and e_1^*, \dots, e_n^* to be a basis of V^* . We then considered multi-indices $I = (i_1, \dots, i_k)$, $1 \leq i_r \leq n$ and defined $\{e_I^* = e_{i_1}^* \otimes \dots \otimes e_{i_k}^*, I \text{ a multi-index}\}$ to be a basis of \mathcal{L}^k . For any multi-index $J = (j_1, \dots, j_k)$, we had

$$e_I^*(e_{j_1}, \dots, e_{j_k}) = \begin{cases} 1 & \text{if } I = J, \\ 0 & \text{if } I \neq J. \end{cases} \quad (4.59)$$

Definition 4.21. A multi-index $I = (i_1, \dots, i_k)$ is *repeating* if $i_r = i_s$ for some $r < s$.

Definition 4.22. The multi-index I is *strictly increasing* if $1 \leq i_1 < \dots < i_k \leq n$.

Notation. Given $\sigma \in S_k$ and $I = (i_1, \dots, i_k)$, we denote $I^\sigma = (i_{\sigma(1)}, \dots, i_{\sigma(k)})$.

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

$$e_J^* = e_{I^\sigma}^* = e_{\sigma(i_1)}^* \otimes \dots \otimes e_{\sigma(i_k)}^* = (e_I^*)^\sigma. \quad (4.60)$$

Define $\psi_I = \text{Alt}(e_I^*)$.

Theorem 4.23. 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

$$\psi_I(e_{j_1}, \dots, e_{j_k}) = \begin{cases} 1 & \text{if } I = J, \\ 0 & \text{if } I \neq J. \end{cases} \quad (4.61)$$

Proof. 1.

$$\begin{aligned} \psi_{I^\sigma} &= \text{Alt } e_{I^\sigma}^* \\ &= \text{Alt } ((e_I^*)^\sigma) \\ &= (-1)^\sigma \text{Alt } e_I^* \\ &= (-1)^\sigma \psi_I. \end{aligned} \quad (4.62)$$

2. Suppose that I is repeating. Then $I = I^\tau$ for some transposition τ . So $\psi_I = (-1)^\tau \psi_I$. But (as you proved in the homework) $(-1)^\tau = -1$, so $\psi_I = 0$.

3.

$$\begin{aligned}\psi_I &= \text{Alt}(e_I^*) \\ &= \sum_{\tau} (-1)^\tau e_{I^\tau}^*,\end{aligned}\tag{4.63}$$

so

$$\psi_I(e_{j_1}, \dots, e_{j_k}) = \sum_{\tau} (-1)^\tau \underbrace{e_{I^\tau}^*(e_{j_1}, \dots, e_{j_k})}_{\begin{cases} 1 & \text{if } I^\tau = J, \\ 0 & \text{if } I^\tau \neq J. \end{cases}}\tag{4.64}$$

But $I^\tau = J$ only if τ is the identity permutation (because both I^τ and J are strictly increasing). The only non-zero term in the sum is when τ is the identity permutation, so

$$\psi_I(e_{j_1}, \dots, e_{j_k}) = \begin{cases} 1 & \text{if } I = J, \\ 0 & \text{if } I \neq J. \end{cases}\tag{4.65}$$

□

Corollary 5. *The alternating k -tensors ψ_I , where I is strictly increasing, are a basis of $\mathcal{A}^k(V)$.*

Proof. Take $T \in \mathcal{A}^k(V)$. The tensor T can be expanded as $T = \sum c_I e_I^*$. So

$$\begin{aligned}\text{Alt}(T) &= k! \sum c_I \text{Alt}(e_I^*) \\ &= k! \sum c_I \psi_I.\end{aligned}\tag{4.66}$$

If I is repeating, then $\psi_I = 0$. If I is non-repeating, then $I = J^\sigma$, where J is strictly increasing. Then $\psi_I = (-1)^\sigma \psi_J$.

So, we can replace all multi-indices in the sum by strictly increasing multi-indices,

$$T = \sum a_I \psi_I, \quad I\text{'s strictly increasing.}\tag{4.67}$$

Therefore, the ψ_I 's span $\mathcal{A}^k(V)$. Moreover, the ψ_I 's are a basis if and only if the a_i 's are unique. We show that the a_I 's are unique.

Let J be any strictly increasing multi-index. Then

$$\begin{aligned}T(e_{j_1}, \dots, e_{j_k}) &= \sum a_I \psi(e_{j_1}, \dots, e_{j_k}) \\ &= a_J,\end{aligned}\tag{4.68}$$

by property (3) of the previous theorem. Therefore, the ψ_I 's are a basis of $\mathcal{A}^k(V)$. □

Lecture 19

We begin with a review of tensors and alternating tensors.

We defined $\mathcal{L}^k(V)$ to be the set of k -linear maps $T : V^k \rightarrow \mathbb{R}$. We defined e_1, \dots, e_n to be a basis of V and e_1^*, \dots, e_n^* to be a basis of V^* . We also defined $\{e_I^* = e_{i_1}^* \otimes \dots \otimes e_{i_k}^*\}$ to be a basis of $\mathcal{L}^k(V)$, where $I = (i_1, \dots, i_k)$, $1 \leq i_r \leq n$ is a multi-index. This showed that $\dim \mathcal{L}^k = n^k$.

We defined the permutation operation on a tensor. For $\sigma \in S_n$ and $T \in \mathcal{L}^k$, we defined $T^\sigma \in \mathcal{L}^k$ by $T^\sigma(v_1, \dots, v_k) = T(v_{\sigma^{-1}(1)}, \dots, v_{\sigma^{-1}(k)})$. Then we defined that T is *alternating* if $T^\sigma = (-1)^\sigma T$. We defined $\mathcal{A}^k = \mathcal{A}^k(V)$ to be the space of all alternating k -tensors.

We defined the alternating operator $\text{Alt} : \mathcal{L}^k \rightarrow \mathcal{A}^k$ by $\text{Alt}(T) = \sum (-1)^\sigma T^\sigma$, and we defined $\psi_I = \text{Alt}(e_I^*)$, where $I = (i_1, \dots, i_k)$ is a strictly increasing multi-index. We proved the following theorem:

Theorem 4.24. *The ψ_I 's (where I is strictly increasing) are a basis for $\mathcal{A}^k(V)$.*

Corollary 6. *If $0 \leq k \leq n$, then*

$$\dim \mathcal{A}^k = \binom{n}{k} = \frac{n!}{k!(n-k)!}. \quad (4.69)$$

Corollary 7. *If $k > n$, then $\mathcal{A}^k = \{0\}$.*

We now ask what is the kernel of Alt ? That is, for which $T \in \mathcal{L}^k$ is $\text{Alt}(T) = 0$?

Let $T \in \mathcal{L}^k$ be a decomposable k -tensor, $T = \ell_1 \otimes \dots \otimes \ell_k$, where each $\ell_i \in V^*$.

Definition 4.25. The k -tensor T is *redundant* if $\ell_i = \ell_{i+1}$ for some $1 \leq i \leq k-1$.

We define

$$\mathcal{I}^k \equiv \text{Span} \{ \text{redundant } k\text{-tensors} \}. \quad (4.70)$$

Claim. *If $T \in \mathcal{I}^k$, then $\text{Alt}(T) = 0$.*

Proof. It suffices to prove this for $T = \ell_1 \otimes \dots \otimes \ell_k$, where $\ell_1 = \ell_{i+1}$ (T is redundant).

Let $\tau = \tau_{i,i+1} \in S_k$. So, $T^\tau = T$. But

$$\begin{aligned} \text{Alt}(T^\tau) &= (-1)^\tau \text{Alt}(T) \\ &= -\text{Alt}(T), \end{aligned} \quad (4.71)$$

so $\text{Alt}(T) = 0$. □

Claim. *Suppose that $T \in \mathcal{I}^k$ and $T' \in \mathcal{L}^m$. Then $T' \otimes T \in \mathcal{I}^{k+m}$ and $T \otimes T' \in \mathcal{I}^{k+m}$.*

Proof. We can assume that T and T' are both decomposable tensors.

$$T = \ell_1 \otimes \cdots \otimes \ell_k, \quad \ell_i = \ell_{i+1}, \quad (4.72)$$

$$T' = \ell'_1 \otimes \cdots \otimes \ell'_m, \quad (4.73)$$

$$T \otimes T' = \ell_1 \otimes \cdots \otimes \underbrace{\ell_i \otimes \ell_{i+1}}_{\text{a redundancy}} \otimes \cdots \otimes \ell_k \otimes \ell'_1 \otimes \cdots \otimes \ell'_m \quad (4.74)$$

$$\in \mathcal{I}^{k+m}. \quad (4.75)$$

A similar argument holds for $T' \otimes T$. \square

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

$$T = (-1)^\sigma T^\sigma + W. \quad (4.76)$$

Proof. In proving this we can assume that T is decomposable. That is, $T = \ell_1 \otimes \cdots \otimes \ell_k$.

We first check the case $k = 2$. Let $T = \ell_1 \otimes \ell_2$. The only (non-identity) permutation is $\sigma = \tau_{1,2}$. In this case, $T = (-1)^\sigma T^\sigma + W$ becomes $W = T + T^\sigma$, so

$$\begin{aligned} W &= T + T^\sigma \\ &= \ell_1 \otimes \ell_2 + \ell_2 \otimes \ell_1 \\ &= (\ell_1 + \ell_2) \otimes (\ell_1 + \ell_2) - \ell_1 \otimes \ell_1 - \ell_2 \otimes \ell_2 \\ &\in \mathcal{I}^2. \end{aligned} \quad (4.77)$$

We now check the case k is arbitrary. Let $T = \ell_1 \otimes \cdots \otimes \ell_k$ and $\sigma = \tau_1 \tau_2 \dots \tau_r \in S_k$, where the τ_i 's are elementary transpositions. We will prove that $W \in \mathcal{I}^k$ by induction on r .

- Case $r = 1$: Then $\sigma = \tau_{i,i+1}$, and

$$\begin{aligned} W &= T + T^\sigma \\ &= (\ell_1 \otimes \cdots \otimes \ell_k) + (\ell_1 \otimes \cdots \otimes \ell_k)^\sigma \\ &= \ell_1 \otimes \cdots \otimes \ell_{i-1} \otimes (\ell_i \otimes \ell_{i+1} + \ell_{i+1} \otimes \ell_i) \otimes \ell_{i+2} \otimes \cdots \otimes \ell_k \\ &\in \mathcal{I}^k, \end{aligned} \quad (4.78)$$

because $(\ell_i \otimes \ell_{i+1} + \ell_{i+1} \otimes \ell_i) \in \mathcal{I}^k$.

- Induction step $((r-1) \implies r)$: Let $\beta = \tau_2 \dots \tau_r$, and let $\tau = \tau_1$ so that $\sigma = \tau_1 \tau_2 \dots \tau_r = \tau \beta$. Then

$$T^\sigma = (T^\beta)^\tau. \quad (4.79)$$

By induction, we know that

$$T^\beta = (-1)^\beta T + W, \quad (4.80)$$

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

$$\begin{aligned} T^\sigma &= (-1)^\beta T^\tau + W^\tau \\ &= (-1)^\beta (-1)^\tau T + W^\tau \\ &= (-1)^\sigma T + W^\tau, \end{aligned} \tag{4.81}$$

where $W^\tau = (-1)^\tau W + W' \in \mathcal{I}^k$.

□

Corollary 8. For every $T \in \mathcal{L}^k$,

$$\text{Alt}(T) = k!T + W \tag{4.82}$$

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

Proof.

$$\text{Alt}(T) = \sum_{\sigma} (-1)^\sigma T^\sigma, \tag{4.83}$$

but we know that $T^\sigma = (-1)^\sigma T + W_\sigma$, for some $W_\sigma \in \mathcal{I}^k$, so

$$\begin{aligned} \text{Alt}(T) &= \sum_{\sigma} (T + (-1)^\sigma W_\sigma) \\ &= k!T + W, \end{aligned} \tag{4.84}$$

where $W = \sum_{\sigma} (-1)^\sigma W_\sigma \in \mathcal{I}^k$.

□

Theorem 4.26. Every $T \in \mathcal{L}^k$ can be written uniquely as a sum

$$T = T_1 + T_2, \tag{4.85}$$

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

Proof. We know that $\text{Alt}(T) = k!T + W$, for some $W \in \mathcal{I}^k$. Solving for T , we get

$$T = \underbrace{\frac{1}{k!} \text{Alt}(T)}_{T_1} - \underbrace{\frac{1}{k!} W}_{T_2}. \tag{4.86}$$

We check uniqueness:

$$\text{Alt}(T) = \underbrace{\text{Alt}(T_1)}_{k!T_1} + \underbrace{\text{Alt}(T_2)}_0, \tag{4.87}$$

so T_1 is unique, which implies that T_2 is also unique.

□

Claim.

$$\mathcal{I}^k = \ker \text{Alt}. \tag{4.88}$$

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

$$T = -\frac{1}{k!}W, \quad W \in \mathcal{I}^k, \quad (4.89)$$

so $T \in \mathcal{I}^k$. □

The space \mathcal{I}^k is a subspace of \mathcal{L}^k , so we can form the quotient space

$$\Lambda^k(V^*) \equiv \mathcal{L}^k / \mathcal{I}^k. \quad (4.90)$$

What's up with this notation $\Lambda^k(V^*)$? We motivate this notation with the case $k = 1$. There are no redundant 1-tensors, so $\mathcal{I}^1 = \{0\}$, and we already know that $\mathcal{L}^1 = V^*$. So

$$\Lambda^1(V^*) = V^* / \mathcal{I}^1 = \mathcal{L}^1 = V^*. \quad (4.91)$$

Define the map $\pi : \mathcal{L}^k \rightarrow \mathcal{L}^k / \mathcal{I}^k$. The map π is onto, and $\ker \pi = \mathcal{I}^k$.

Claim. *The map π maps \mathcal{A}^k bijectively onto $\Lambda^k(V^*)$.*

Proof. Every element of Λ^k is of the form $\pi(T)$ for some $T \in \mathcal{L}^k$. We can write $T = T_1 + T_2$, where $T_1 \in \mathcal{A}^k$ and $T_2 \in \mathcal{I}^k$. So,

$$\begin{aligned} \pi(T) &= \pi(T_1) + \pi(T_2) \\ &= \pi(T_1) + 0 \\ &= \pi(T_1). \end{aligned} \quad (4.92)$$

So, π maps \mathcal{A}^k onto Λ^k . Now we show that π is one-to-one. If $T \in \mathcal{A}^k$ and $\pi(T) = 0$, then $T \in \mathcal{I}^k$ as well. We know that $\mathcal{A}^k \cap \mathcal{I}^k = \{0\}$, so π is bijective. □

We have shown that

$$\mathcal{A}^k(V) \cong \Lambda^k(V^*). \quad (4.93)$$

The space $\Lambda^k(V^*)$ is not mentioned in Munkres, but sometimes it is useful to look at the same space in two different ways.

Lecture 20

We begin with a review of last lecture.

Consider a vector space V . A tensor $T \in \mathcal{L}^k$ is *decomposable* if $T = \ell_1 \otimes \cdots \otimes \ell_k$, $\ell_i \in \mathcal{L}^1 = V^*$. A decomposable tensor T is *redundant* if $\ell_i = \ell_{i+1}$ for some i . We define

$$\mathcal{I}^k = \mathcal{I}^k(V) = \text{Span} \{ \text{redundant } k\text{-tensors} \}. \quad (4.94)$$

Because $\mathcal{I}^k \subseteq \mathcal{L}^k$, we can take the quotient space

$$\Lambda^k = \Lambda^k(V^*) = \mathcal{L}^k / \mathcal{I}^k, \quad (4.95)$$

defining the map

$$\pi : \mathcal{L}^k \rightarrow \Lambda^k. \quad (4.96)$$

We denote by $\mathcal{A}^k(V)$ the set of all alternating k -tensors. We repeat the main theorem from last lecture:

Theorem 4.27. *The map π maps \mathcal{A}^k bijectively onto Λ^k . So, $\mathcal{A}^k \cong \Lambda^k$.*

It is easier to understand the space \mathcal{A}^k , but many theorems are much simpler when using Λ^k . This ends the review of last lecture.

4.6 Wedge Product

Now, let $T_1 \in \mathcal{I}^{k_1}$ and $T_2 \in \mathcal{L}^{k_2}$. Then $T_1 \otimes T_2$ and $T_2 \otimes T_1$ are in \mathcal{I}^k , where $k = k_1 + k_2$. The following is an example of the usefulness of Λ^k .

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

$$\pi(T_1 \otimes T_2) = \mu_1 \wedge \mu_2 \in \Lambda^k. \quad (4.97)$$

Claim. *The product $\mu_i \wedge \mu_2$ is well-defined.*

Proof. Take any tensors $T'_i \in \mathcal{L}^{k_i}$ with $\pi(T'_i) = \mu_i$. We check that

$$\pi(T'_1 \otimes T'_2) = \pi(T_1 \otimes T_2). \quad (4.98)$$

We can write

$$T'_1 = T_1 + W_1, \text{ where } W_1 \in \mathcal{I}^{k_1}, \quad (4.99)$$

$$T'_2 = T_2 + W_2, \text{ where } W_2 \in \mathcal{I}^{k_2}. \quad (4.100)$$

Then,

$$T'_1 \otimes T'_2 = T_1 \otimes T_2 + \underbrace{W_1 \otimes T_2 + T_1 \otimes W_2 + W_1 \otimes W_2}_{\in \mathcal{I}^k}, \quad (4.101)$$

so

$$\mu_1 \wedge \mu_2 \equiv \pi(T'_1 \otimes T'_2) = \pi(T_1 \otimes T_2). \quad (4.102)$$

□

This product (\wedge) is called the *wedge product*. We can define higher order wedge products. Given $\mu_i \in \Lambda^{k_i}$, $i = 1, 2, 3$, where $\mu = \pi(T_i)$, we define

$$\mu_1 \wedge \mu_2 \wedge \mu_3 = \pi(T_1 \otimes T_2 \otimes T_3). \quad (4.103)$$

We leave as an exercise to show the following claim.

Claim.

$$\begin{aligned} \mu_1 \wedge \mu_2 \wedge \mu_3 &= (\mu_1 \wedge \mu_2) \wedge \mu_3 \\ &= \mu_1 \wedge (\mu_2 \wedge \mu_3). \end{aligned} \quad (4.104)$$

Proof Hint: This triple product law also holds for the tensor product. \square

We leave as an exercise to show that the two distributive laws hold:

Claim. *If $k_1 = k_2$, then*

$$(\mu_1 + \mu_2) \wedge \mu_3 = \mu_1 \wedge \mu_3 + \mu_2 \wedge \mu_3. \quad (4.105)$$

If $k_2 = k_3$, then

$$\mu_1 \wedge (\mu_2 + \mu_3) = \mu_1 \wedge \mu_2 + \mu_1 \wedge \mu_3. \quad (4.106)$$

Remember that $\mathcal{I}^1 = \{0\}$, so $\Lambda^1 = \Lambda^1/\mathcal{I}^1 = \mathcal{L}^1 = \mathcal{L}^1(V) = V^*$. That is, $\Lambda^1(V^*) = V^*$.

Definition 4.28. The element $\mu \in \Lambda^k$ is *decomposable* if it is of the form $\mu = \ell_1 \wedge \cdots \wedge \ell_k$, where each $\ell_i \in \Lambda^1 = V^*$.

That means that $\mu = \pi(\ell_1 \otimes \cdots \otimes \ell_k)$ is the projection of a decomposable k -tensor.

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

Definition 4.29.

$$\omega^\sigma = \pi(T^\sigma). \quad (4.107)$$

We need to check that this definition does not depend on the choice of T .

Claim. *Define $\omega^\sigma = \pi(T^\sigma)$. Then,*

1. *The above definition does not depend on the choice of T ,*
2. $\omega^\sigma = (-1)^\sigma \omega$.

Proof. 1. Last lecture we proved that for $T \in \mathcal{L}^k$,

$$T^\sigma = (-1)^\sigma T + W, \quad (4.108)$$

for some $W \in \mathcal{I}^k$. Hence, if $T \in \mathcal{I}^k$, then $T^\sigma \in \mathcal{I}^k$. If $\omega = \pi(T) = \pi(T')$, then $T' - T \in \mathcal{I}^k$. Thus, $(T')^\sigma - T^\sigma \in \mathcal{I}^k$, so $\omega^\sigma = \pi((T')^\sigma) = \pi(T^\sigma)$.

2.

$$T^\sigma = (-1)^\sigma T + W, \quad (4.109)$$

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

$$\pi(T^\sigma) = (-1)^\sigma \pi(T). \quad (4.110)$$

That is,

$$\omega^\sigma = (-1)^\sigma \omega. \quad (4.111)$$

□

Suppose ω is decomposable, so $\omega = \ell_1 \wedge \cdots \wedge \ell_k$, $\ell_i \in V^*$. Then $\omega = \pi(\ell_1 \wedge \cdots \wedge \ell_k)$, so

$$\begin{aligned} \omega^\sigma &= \pi((\ell_1 \otimes \cdots \otimes \ell_k)^\sigma) \\ &= \pi(\ell_{\sigma(1)} \otimes \cdots \otimes \ell_{\sigma(k)}) \\ &= \ell_{\sigma(1)} \wedge \cdots \wedge \ell_{\sigma(k)}. \end{aligned} \quad (4.112)$$

Using the previous claim,

$$\ell_{\sigma(1)} \wedge \cdots \wedge \ell_{\sigma(k)} = (-1)^\sigma \ell_1 \wedge \cdots \wedge \ell_k. \quad (4.113)$$

For example, if $k = 2$, then $\sigma = \tau_{1,2}$. So, $\ell_2 \wedge \ell_1 = -\ell_1 \wedge \ell_2$. In the case $k = 3$, we find that

$$\begin{aligned} (\ell_1 \wedge \ell_2) \wedge \ell_3 &= \ell_1 \wedge (\ell_2 \wedge \ell_3) \\ &= -\ell_1 \wedge (\ell_3 \wedge \ell_2) = -(\ell_1 \wedge \ell_3) \wedge \ell_2 \\ &= \ell_3 \wedge (\ell_1 \wedge \ell_2). \end{aligned} \quad (4.114)$$

This motivates the following claim, the proof of which we leave as an exercise.

Claim. *If $\mu \in \Lambda^2$ and $\ell \in \Lambda^1$, then*

$$\mu \wedge \ell = \ell \wedge \mu. \quad (4.115)$$

Proof Hint: Write out μ as a linear combination of decomposable elements of Λ^2 . □

Now, suppose $k = 4$. Moving ℓ_3 and ℓ_4 the same distance, we find that

$$(\ell_1 \wedge \ell_2) \wedge (\ell_3 \wedge \ell_4) = (\ell_3 \wedge \ell_4) \wedge (\ell_1 \wedge \ell_2). \quad (4.116)$$

The proof of the following is an exercise.

Claim. *If $\mu \in \Lambda^2$ and $\nu \in \Lambda^2$, then*

$$\mu \wedge \nu = \nu \wedge \mu. \quad (4.117)$$

We generalize the above claims in the following:

Claim. Left $\mu \in \Lambda^k$ and $\nu \in \Lambda^\ell$. Then

$$\mu \wedge \nu = (-1)^{k\ell} \nu \wedge \mu. \quad (4.118)$$

Proof Hint: First assume k is even, and write out μ as a product of elements all of degree two. Second, assume that k is odd. \square

Now we try to find a basis for $\Lambda^k(V^*)$. We begin with

$$e_1, \dots, e_n \text{ a basis of } V, \quad (4.119)$$

$$e_1^*, \dots, e_n^* \text{ a basis of } V^*, \quad (4.120)$$

$$e_I^* = e_{i_1}^* \otimes \cdots \otimes e_{i_k}^*, \quad I = (i_1, \dots, i_k), \quad 1 \leq i_r \leq n, \text{ a basis of } \mathcal{L}^k, \quad (4.121)$$

$$\psi_I = \text{Alt}(e_I^*), \quad I \text{'s strictly increasing, a basis of } \mathcal{A}^k(V). \quad (4.122)$$

We know that π maps \mathcal{A}^k bijectively onto Λ^k , so $\pi(\psi_I)$, where I is strictly increasing, are a basis of $\Lambda^k(V^*)$.

$$\psi_I = \text{Alt } e_I^* = \sum (-1)^\sigma (e_I^*)^\sigma. \quad (4.123)$$

So,

$$\begin{aligned} \pi(\psi_I) &= \sum (-1)^\sigma \pi((e_I^*)^\sigma) \\ &= \sum (-1)^\sigma (-1)^\sigma \pi(e_I^*) \\ &= k! \pi(e_I^*) \\ &\equiv k! \tilde{e}_I. \end{aligned} \quad (4.124)$$

Theorem 4.30. The elements of $\Lambda^k(V^*)$

$$\tilde{e}_{i_1}^* \wedge \cdots \wedge \tilde{e}_{i_k}^*, \quad 1 \leq i_1 < \dots < i_k \leq n \quad (4.125)$$

are a basis of $\Lambda^k(V^*)$.

Proof. The proof is above. \square

Let V, W be vector spaces, and let $A : V \rightarrow W$ be a linear map. We previously defined the pullback operator $A^* : \mathcal{L}^k(W) \rightarrow \mathcal{L}^k(V)$. Also, given $T_i \in \mathcal{L}^{k_i}(W)$, $i = 1, 2$, we showed that $A^*(T_1 \otimes T_2) = A^*T_1 \otimes A^*T_2$. So, if $T = \ell_1 \otimes \cdots \otimes \ell_k \in \mathcal{L}_k(W)$ is decomposable, then

$$A^*T = A^*\ell_1 \otimes \cdots \otimes A^*\ell_k, \quad \ell_i \in W^*. \quad (4.126)$$

If $\ell_i = \ell_{i+1}$, then $A^*\ell_i = A^*\ell_{i+1}$. This shows that if $\ell_1 \otimes \cdots \otimes \ell_k$ is redundant, then $A^*(\ell_1 \otimes \cdots \otimes \ell_k)$ is also redundant. So,

$$A^*\mathcal{I}^k(W) \subseteq \mathcal{I}^k(V). \quad (4.127)$$

Let $\mu \in \Lambda^k(W^*)$, so $\mu = \pi(T)$ for some $T \in \mathcal{L}^k(W)$. We can pullback to get $\pi(A^*T) \in \Lambda^k(V^*)$.

Definition 4.31. $A^*\mu = \pi(A^*T)$.

This definition makes sense. If $\mu = \pi(T) = \pi(T')$, then $T' - T \in \mathcal{I}^k(W)$. So $A^*T' - A^*T \in \mathcal{I}^k(V)$, which shows that $A^*\mu = \pi(A^*T') = \pi(A^*T)$.

We ask in the homework for you to show that the pullback operation is linear and that

$$A^*(\mu_1 \wedge \mu_2) = A^*\mu_1 \wedge A^*\mu_2. \quad (4.128)$$

Lecture 21

Let V, W be vector spaces, and let $A : V \rightarrow W$ be a linear map. We defined the pullback operation $A^* : W^* \rightarrow V^*$. Last time we defined another pullback operator having the form $A^* : \Lambda^k(W^*) \rightarrow \Lambda^k(V^*)$. This new pullback operator has the following properties:

1. A^* is linear.
2. If $\omega_i \in \Lambda^{k_i}(W^*)$, $i = 1, 2$, then $A^*\omega_1 \wedge \omega_2 = A^*\omega_1 \wedge \omega_2$.
3. If ω is decomposable, that is if $\omega = \ell_1 \wedge \cdots \wedge \ell_k$ where $\ell_i \in W^*$, then $A^*\omega = A^*\ell_1 \wedge \cdots \wedge A^*\ell_k$.
4. Suppose that U is a vector space and that $B : W \rightarrow U$ is a linear map. Then, for every $\omega \in \Lambda^k(U^*)$, $A^*B^*\omega = (BA)^*\omega$.

4.7 Determinant

Today we focus on the pullback operation in the special case where $\dim V = n$. So, we are studying $\Lambda^n(V^*)$, which is called the *nth exterior power of V* .

Note that $\dim \Lambda^n(V^*) = 1$.

Given a linear map $A : V \rightarrow V$, what is the pullback operator

$$A^* : \Lambda^n(V^*) \rightarrow \Lambda^n(V^*)? \quad (4.129)$$

Since it is a linear map from a one dimensional vector space to a one dimensional vector space, the pullback operator A^* is simply multiplication by some constant λ_A . That is, for all $\omega \in \Lambda^n(V^*)$, $A^*\omega = \lambda_A\omega$.

Definition 4.32. The *determinant of A* is

$$\det(A) = \lambda_A. \quad (4.130)$$

The determinant has the following properties:

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

Proof: Let $\omega \in \Lambda^n(V^*)$. Then

$$\begin{aligned} (AB)^*\omega &= \det(AB)\omega \\ &= B^*(A^*\omega) \\ &= B^*(\det A)\omega \\ &= \det(A) \det(B)\omega. \end{aligned} \quad (4.131)$$

3. If A is onto, then $\det(A) \neq 0$.

Proof: Suppose that $A : V \rightarrow V$ is onto. Then there exists an inverse linear map $A^{-1} : V \rightarrow V$ such that $AA^{-1} = I$. So, $\det(A)\det(A^{-1}) = 1$.

4. If A is not onto, then $\det(A) = 0$.

Proof: Let $W = \text{Im}(A)$. If A is not onto, then $\dim W < \dim V$. Let $B : V \rightarrow W$ be the map A regarded as a map of V into W , and let $\iota_W : W \rightarrow V$ be inclusion. So, $A = \iota_W B$. For all $\omega \in \Lambda^n(V^*)$, $A^*\omega = B^*\iota_W^*\omega$. Note that $\iota_W^*\omega \in \Lambda^n(W^*) = \{0\}$ because $\dim W < n$. So, $A^*\omega = B^*\iota_W^*\omega = 0$, which shows that $\det(A) = 0$.

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

$$e_1, \dots, e_n \text{ basis of } V, \quad (4.132)$$

$$e_1^*, \dots, e_n^* \text{ dual basis of } V^*, \quad (4.133)$$

$$f_1, \dots, f_n \text{ basis of } W, \quad (4.134)$$

$$f_1^*, \dots, f_n^* \text{ dual basis of } W^*. \quad (4.135)$$

We can write $Ae_i = \sum a_{ij}f_j$, so that A has the associated matrix $A \sim [a_{ij}]$. Then $A^*f_j^* = \sum a_{jk}e_k^*$. Take $\omega = f_1^* \wedge \dots \wedge f_n^* \in \Lambda^n(W^*)$, which is a basis vector of $\Lambda^n(W^*)$. Let us compute its pullback:

$$\begin{aligned} A^*(f_1^* \wedge \dots \wedge f_n^*) &= \left(\sum_{k_1=1}^n a_{1,k_1} e_{k_1}^* \right) \wedge \dots \wedge \left(\sum_{k_n=1}^n a_{n,k_n} e_{k_n}^* \right) \\ &= \sum_{k_1, \dots, k_n} (a_{1,k_1} \dots a_{n,k_n}) e_{k_1}^* \wedge \dots \wedge e_{k_n}^*. \end{aligned} \quad (4.136)$$

Note that if $k_r = k_s$, where $r \neq s$, then $e_{k_1}^* \wedge \dots \wedge e_{k_n}^* = 0$. If there are no repetitions, then there exists $\sigma \in S_n$ such that $k_i = \sigma(i)$. Thus,

$$\begin{aligned} A^*(f_1^* \wedge \dots \wedge f_n^*) &= \sum_{\sigma} a_{1,\sigma(1)} \dots a_{n,\sigma(n)} e_{\sigma(1)}^* \wedge \dots \wedge e_{\sigma(n)}^* \\ &= \left(\sum_{\sigma} (-1)^{\sigma} a_{1,\sigma(1)} \dots a_{n,\sigma(n)} \right) e_1^* \wedge \dots \wedge e_n^*. \end{aligned} \quad (4.137)$$

Therefore,

$$\det[a_{ij}] = \sum_{\sigma} (-1)^{\sigma} a_{1,\sigma(1)} \dots a_{n,\sigma(n)}. \quad (4.138)$$

In the case where $W = V$ and each $e_i = f_i$, we set $\omega = e_1^* \wedge \dots \wedge e_n^*$, and we get $A^*\omega = \det[a_{ij}]\omega$. So, $\det(A) = \det[a_{ij}]$.

For basic facts about determinants, see Munkres section 2. We will use these results quite a lot in future lectures. We list some of the basic results below.

Let $A = [a_{ij}]$ be an $n \times n$ matrix.

1. $\det(A) = \det(A^t)$. You should prove this as an exercise. You should explain the following steps:

$$\begin{aligned} \det(A) &= \sum_{\sigma} (-1)^{\sigma} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \\ &= \sum_{\tau} (-1)^{\tau} a_{\tau(1),1} \cdots a_{\tau(n),n}, \text{ where } \tau = \sigma^{-1} \\ &= \det(A^t). \end{aligned} \tag{4.139}$$

2. Let

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}, \tag{4.140}$$

where B is $k \times k$, C is $k \times \ell$, D is $\ell \times \ell$, and $n = k + \ell$. Then

$$\det(A) = \det(B) \det(D) \tag{4.141}$$

4.8 Orientations of Vector Spaces

Let $\ell \subseteq \mathbb{R}^2$ be a line through the origin. Then $\ell - \{0\}$ has two connected components. An *orientation* of ℓ is a choice of one of these components.

More generally, given a one-dimensional vector space \mathbb{L} , the set $\mathbb{L} - \{0\}$ has two connected components. Choose $v \in \mathbb{L} - \{0\}$. Then the two components are

$$\{\lambda v : \lambda \in \mathbb{R}_+\} \text{ and } \{-\lambda v : \lambda \in \mathbb{R}_+\}. \tag{4.142}$$

Definition 4.33. An *orientation* of \mathbb{L} is a choice of one of these components, usually labeled \mathbb{L}_+ . We define

$$v \in \mathbb{L}_+ \iff v \text{ is positively oriented.} \tag{4.143}$$

Let V be an n -dimensional vector space. Then $\Lambda^n(V^*)$ is a 1-dimensional vector space.

Definition 4.34. An *orientation* of V is an orientation of $\Lambda^n(V^*)$. That is, a choice of $\Lambda^n(V^*)_+$.

Suppose e_1, \dots, e_n is a basis of V , so e_1^*, \dots, e_n^* is the dual basis of V^* . Let $\omega = e_1^* \wedge \cdots \wedge e_n^* \in \Lambda^n(V^*) - \{0\}$.

Definition 4.35. The basis e_1, \dots, e_n is *positively oriented* if $\omega \in \Lambda^n(V^*)_+$.

Let f_1, \dots, f_n be another basis of V and f_1^*, \dots, f_n^* its dual basis. Let $w' = f_1^* \wedge \cdots \wedge f_n^*$. We ask: How is w' related to ω ? The answer: If $f_j = \sum a_{ij} e_i$, then $w' = \det[a_{ij}] \omega$. So, if e_1, \dots, e_n is positively oriented, then f_1, \dots, f_n is positively oriented if and only if $\det[a_{ij}] > 0$.

Suppose V is an n -dimensional vector space and that W is a k -dimensional subspace of V .

Claim. *If V and V/W are given orientations, then W acquires from these orientations a natural subspace orientation.*

Idea of proof: Let $\pi : V \rightarrow V/W$ be the canonical map, and choose a basis e_1, \dots, e_n of V such that $e_{\ell+1}, \dots, e_n$ is a basis of W and such that $\pi(e_1), \dots, \pi(e_\ell)$ is a basis of V/W , where $\ell = n - k$.

Replacing e_1 by $-e_1$ if necessary, we can assume that $\pi(e_1), \dots, \pi(e_\ell)$ is an oriented basis of V/W . Replacing e_n by $-e_n$ if necessary, we can assume that e_1, \dots, e_n is an oriented basis of V . Now, give W the orientation for which $e_{\ell+1}, \dots, e_n$ is an oriented basis of W . One should check that this choice of orientation for W is independent of the choice of basis (this is explained in the Multi-linear Algebra notes). \square

Lecture 22

In \mathbb{R}^3 we had the operators grad, div, and curl. What are the analogues in \mathbb{R}^n ? Answering this question is the goal of today's lecture.

4.9 Tangent Spaces and k -forms

Let $p \in \mathbb{R}^n$.

Definition 4.36. The *tangent space of p in \mathbb{R}^n* is

$$T_p\mathbb{R}^n = \{(p, v) : v \in \mathbb{R}^n\}. \quad (4.144)$$

We identify the tangent space with \mathbb{R}^n via the identification

$$T_p\mathbb{R}^n \cong \mathbb{R}^n \quad (4.145)$$

$$(p, v) \rightarrow v. \quad (4.146)$$

Via this identification, the vector space structure on \mathbb{R}^n gives a vector space structure on $T_p\mathbb{R}^n$.

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

$$df_p : T_p\mathbb{R}^n \rightarrow T_q\mathbb{R}^m \quad (4.147)$$

according to the following diagram:

$$\begin{array}{ccc} T_p\mathbb{R}^n & \xrightarrow{df_p} & T_q\mathbb{R}^m \\ \cong \downarrow & & \cong \uparrow \\ \mathbb{R}^n & \xrightarrow{Df(p)} & \mathbb{R}^m. \end{array} \quad (4.148)$$

So,

$$df_p(p, v) = (q, Df(p)v). \quad (4.149)$$

Definition 4.37. The *cotangent space of \mathbb{R}^n at p* is the space

$$T_p^*\mathbb{R}^n \equiv (T_p\mathbb{R}^n)^*, \quad (4.150)$$

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

Definition 4.38. Let U be an open subset of \mathbb{R}^n . A *k -form on U* is a function ω which assigns to every point $p \in U$ an element ω_p of $\Lambda^k(T_p^*\mathbb{R}^n)$ (the k th exterior power of $T_p^*\mathbb{R}^n$).

Let us look at a simple example. Let $f \in \mathcal{C}^\infty(U)$, $p \in U$, and $c = f(p)$. The map

$$df_p : T_p\mathbb{R}^n \rightarrow T_c\mathbb{R} = \mathbb{R} \quad (4.151)$$

is a linear map of $T_p\mathbb{R}^n$ into \mathbb{R} . That is, $df_p \in T_p^*\mathbb{R}^n$. So, df is the one-form on U which assigns to every $p \in U$ the linear map

$$df_p \in T_p^*\mathbb{R}^n = \Lambda^1(T_p^*\mathbb{R}^n). \quad (4.152)$$

As a second example, let $f, g \in \mathcal{C}^\infty(U)$. Then gdf is the one-form that maps

$$p \in U \rightarrow g(p)df_p \in \Lambda^1(T_p^*\mathbb{R}^n). \quad (4.153)$$

As a third example, let $f, g \in \mathcal{C}^\infty(U)$. Then $\omega = df \wedge dg$ is the two-form that maps

$$p \in U \rightarrow df_p \wedge dg_p. \quad (4.154)$$

Note that $df_p, dg_p \in T_p^*\mathbb{R}^n$, so $df_p \wedge dg_p \in \Lambda^2(T_p^*\mathbb{R}^n)$.

As a fourth and final example, let $f_1, \dots, f_k \in \mathcal{C}^\infty(U)$. Then $df_1 \wedge \dots \wedge df_k$ is the k -form that maps

$$p \in U \rightarrow (df_1)_p \wedge \dots \wedge (df_k)_p. \quad (4.155)$$

Note that each $(df_i)_p \in T_p^*\mathbb{R}^n$, so $(df_1)_p \wedge \dots \wedge (df_k)_p \in \Lambda^k(T_p^*\mathbb{R}^n)$.

Let us now look at what k -forms look like in coordinates. Let e_1, \dots, e_n be the standard basis of \mathbb{R}^n . Let $p \in U$ and let $v_i = (p, e_i)$ for each i . Then, the vectors v_1, \dots, v_n are a basis of $T_p\mathbb{R}^n$.

Suppose we have a map $f \in \mathcal{C}^\infty(U)$. What is $df_p(v_i)$?

$$df_p(v_i) = De_i f(p) = \frac{\partial f}{\partial x_i}(p). \quad (4.156)$$

In particular, letting x_i be the i th coordinate function,

$$(dx_i)_p(v_j) = \frac{\partial x_i}{\partial x_j} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \quad (4.157)$$

So, $(dx_1)_p, \dots, (dx_n)_p$ is the basis of $T_p^*\mathbb{R}^n$ dual to v_1, \dots, v_n .

For any $f \in \mathcal{C}^\infty(U)$,

$$\begin{aligned} df_p(v_j) &= \frac{\partial f}{\partial x_j}(p) \\ &= \left(\sum_i \frac{\partial f}{\partial x_i}(p) (dx_i)_p \right) (v_j) \\ \implies df_p &= \sum \frac{\partial f}{\partial x_i}(p) (dx_i)_p \\ \implies df &= \sum \frac{\partial f}{\partial x_i} dx_i. \end{aligned} \quad (4.158)$$

Since $(dx_1)_p, \dots, (dx_n)_p$ is a basis of $T_p^*\mathbb{R}^n$, the wedge products

$$(dx_I)_p = (dx_{i_1})_p \wedge \cdots \wedge (dx_{i_k})_p, \quad 1 \leq i_1 < \cdots < i_k \leq n, \quad (4.159)$$

(I strictly increasing) are a basis of $\Lambda^k(T_p^*\mathbb{R}^n)$.

Therefore, any element w_p of $\Lambda^k(T_p^*\mathbb{R}^n)$ can be written uniquely as a sum

$$\omega_p = \sum a_I(p)(dx_I)_p, \quad a_I(p) \in \mathbb{R}, \quad (4.160)$$

where the I 's are strictly increasing. Hence, any k -form can be written uniquely as a sum

$$\omega = \sum a_I dx_I, \quad I \text{ strictly increasing}, \quad (4.161)$$

where each a_I is a real-valued function on U . That is, $a_I : U \rightarrow \mathbb{R}$.

Definition 4.39. The k -form ω is $\mathcal{C}^r(U)$ if each $a_I \in \mathcal{C}^r(U)$.

Just to simplify our discussion, from now on we will always take k -forms that are \mathcal{C}^∞ .

Definition 4.40. We define

$$\Omega^k(U) = \text{the set of all } \mathcal{C}^\infty \text{ } k\text{-forms.} \quad (4.162)$$

So, $\omega \in \Omega^k(U)$ implies that $\omega = \sum a_I dx_I$, where $a_I \in \mathcal{C}^\infty(U)$.

Let us now study some basic operations on k -forms.

1. Let $\omega \in \Omega^k(U)$ and let $f \in \mathcal{C}^\infty(U)$. Then $f\omega \in \Omega^k(U)$ is the k -form that maps

$$p \in U \rightarrow f(p)\omega_p \in \Lambda^k(T_p^*\mathbb{R}^n). \quad (4.163)$$

2. Let $\omega_i \in \Omega^k(U)$, $i = 1, 2$. Then $\omega_1 + \omega_2$ is the k -form that maps

$$p \in U \rightarrow (\omega_1)_p + (\omega_2)_p \in \Lambda^k(T_p^*\mathbb{R}^n). \quad (4.164)$$

3. Let $\omega_i \in \Omega^{k_i}(U)$, $i = 1, 2$, and $k = k_1 + k_2$. Then $\omega_1 \wedge \omega_2 \in \Omega^k(U)$ is the k -form that maps

$$p \in U \rightarrow (\omega_1)_p \wedge (\omega_2)_p \in \Lambda^k(T_p^*\mathbb{R}^n), \quad (4.165)$$

since $(\omega_i)_p \in \Lambda^{k_i}(T_p^*\mathbb{R}^n)$.

Definition 4.41. We find it convenient to define $\Lambda^0(T_p^*\mathbb{R}^n) = \mathbb{R}$.

A zero-form f on U is just a real-valued function, so $\Omega^0(U) = \mathcal{C}^\infty(\mathbb{R})$.

Take $f \in \mathcal{C}^\infty(U)$ and $df \in \Omega^1(U)$. This gives an operation

$$d : \Omega^0(U) \rightarrow \Omega^1(U), \quad (4.166)$$

$$f \rightarrow df. \quad (4.167)$$

Let $f, g \in \mathcal{C}^\infty(U)$ (that is, take f, g to be zero-forms). Then $d(fg) = gdf + fdg$. We can think of this operation as a slightly different notation for the gradient operation.

The maps $d : \Omega^k(U) \rightarrow \Omega^{k+1}(U)$, $k = 0, \dots, (n-1)$ give n vector calculus operations.

If $\omega \in \Omega^k(U)$, then ω can be written uniquely as the sum

$$\omega = \sum a_I dx_I, \quad I \text{ strictly increasing}, \quad (4.168)$$

where each $a_I \in \mathcal{C}^\infty(U)$. We define

$$d\omega = \sum da_I \wedge dx_I. \quad (4.169)$$

This operator is the unique operator with the following properties:

1. For $k = 0$, this is the operation we already defined, $df = \sum \frac{\partial f}{\partial x_i} dx_i$.
2. If $\omega \in \Omega^k$, then $d(d\omega) = 0$.
3. If $\omega_i \in \Omega^{k_i}(U)$, $i = 1, 2$, then $d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^{k_1} \omega_1 \wedge d\omega_2$.

Let $a \in \mathcal{C}^\infty(U)$, and $adx_I \in \Omega^k(U)$, I strictly increasing. Then

$$d(adx_I) = da \wedge dx_I. \quad (4.170)$$

Suppose that I is not strictly increasing. Then

$$\begin{aligned} dx_I &= dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\ &= 0 \quad \text{if } i_r = i_s. \end{aligned} \quad (4.171)$$

If there are no repetitions, then there exists $\sigma \in S_k$ such that $J = I^\sigma$ is strictly increasing. Then

$$dx_J = (-1)^\sigma dx_I, \quad (4.172)$$

so

$$\begin{aligned} d(adx_I) &= (-1)^\sigma d(adx_J) \\ &= (-1)^\sigma da \wedge dx_J \\ &= da \wedge dx_I. \end{aligned} \quad (4.173)$$

Putting this altogether, for *every* multi-index I of length k ,

$$d(adx_I) = da \wedge dx_I. \quad (4.174)$$

Lecture 23

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

$$d : \Omega^k(U) \rightarrow \Omega^{k+1}(U). \quad (4.175)$$

These maps are the n basic vector calculus operations in n -dimensional calculus. We review how d is defined.

For $k = 0$, $\Omega^0(U) = \mathcal{C}^\infty(U)$. Let $f \in \mathcal{C}^\infty(U)$, and let $c = f(p)$, where $p \in U$. The mapping $df_p : T_p\mathbb{R}^n \rightarrow T_c\mathbb{R} = \mathbb{R}$ maps $T_p\mathbb{R}^n$ to \mathbb{R} , so $df_p \in T_p^*\mathbb{R}^n$. The map $df \in \Omega^1(U)$ is a one-form that maps $p \in U$ to $df_p \in T_p^*\mathbb{R}^n$. A formula for this in coordinates is

$$df = \sum \frac{\partial f}{\partial x_i} dx_i. \quad (4.176)$$

In k dimensions, d is a map

$$d : \Omega^k(U) \rightarrow \Omega^{k+1}(U). \quad (4.177)$$

Given $\omega \in \Omega^k(U)$, ω can be written uniquely as

$$\begin{aligned} \omega &= \sum_I a_I dx_I \\ &= \sum_I a_I dx_{i_1} \wedge \cdots \wedge dx_{i_k}, \end{aligned} \quad (4.178)$$

where $i_1 < \cdots < i_k$ and each $a_I \in \mathcal{C}^\infty(U)$. Then, we define

$$\begin{aligned} d\omega &= \sum da_I \wedge dx_I \\ &= \sum_{i,I} \frac{\partial a_I}{\partial x_i} dx_i \wedge dx_I, \end{aligned} \quad (4.179)$$

where each I is strictly increasing.

The following are some basic properties of the differential operator d :

1. If $\mu \in \Omega^k(U)$ and $\nu \in \Omega^\ell(U)$, then

$$d\mu \wedge \nu = d\mu \wedge \nu + (-1)^k \mu \wedge d\nu. \quad (4.180)$$

2. For and $\omega \in \Omega^k(U)$,

$$d(d\omega) = 0. \quad (4.181)$$

Remark. Let I be any multi-index, and let $a_I \in \mathcal{C}^\infty(U)$. Then

$$d(a_I dx_I) = da_I \wedge dx_I. \quad (4.182)$$

We now prove the above two basic properties of the differential operator.

Claim. If $\mu \in \Omega^k(U)$ and $\nu \in \Omega^\ell(U)$, then

$$d\mu \wedge \nu = d\mu \wedge \nu + (-1)^k \mu \wedge d\nu. \quad (4.183)$$

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

$$\mu \wedge \nu = \sum a_I b_J \underbrace{dx_I \wedge dx_J}_{\text{no longer increasing}}. \quad (4.184)$$

Then

$$\begin{aligned} d(\mu \wedge \nu) &= \sum_{i,I,J} \frac{\partial a_I b_J}{\partial x_i} dx_i \wedge dx_I \wedge dx_J \\ &= \sum \frac{\partial a_I}{\partial x_i} b_J dx_i \wedge dx_I \wedge dx_J \quad (I) \\ &\quad + \sum a_I \frac{\partial b_J}{\partial x_i} dx_i \wedge dx_I \wedge dx_J, \quad (II) \end{aligned} \quad (4.185)$$

We calculate sums (I) and (II) separately.

$$\begin{aligned} (I) &= \sum_{i,I,J} \frac{\partial a_I}{\partial x_i} dx_i \wedge dx_I \wedge b_J dx_J \\ &= \left(\sum_{i,I} \frac{\partial a_I}{\partial x_i} dx_i \wedge dx_I \right) \wedge \sum_J b_J dx_J \\ &= d\mu \wedge \nu. \end{aligned} \quad (4.186)$$

$$\begin{aligned} (II) &= \sum_{i,I,J} a_I \frac{\partial b_J}{\partial x_i} dx_i \wedge dx_I \wedge dx_J \\ &= (-1)^k \sum_{i,I,J} a_I dx_I \wedge \frac{\partial b_J}{\partial x_i} dx_i \wedge dx_J \\ &= (-1)^k \left(\sum_I a_I dx_I \right) \wedge \sum_{i,J} \frac{\partial b_J}{\partial x_i} dx_i \wedge dx_J \\ &= (-1)^k \mu \wedge d\nu. \end{aligned} \quad (4.187)$$

So,

$$\begin{aligned} d(\mu \wedge \nu) &= (I) + (II) \\ &= d\mu \wedge \nu + (-1)^k \mu \wedge d\nu. \end{aligned} \quad (4.188)$$

□

Claim. For and $\omega \in \Omega^k(U)$,

$$d(d\omega) = 0. \quad (4.189)$$

Proof. Let $\omega = \sum a_I dx_I$, so

$$d\omega = \sum_{j,I} \frac{\partial a_I}{\partial x_j} dx_j \wedge dx_I. \quad (4.190)$$

Then,

$$d(d\omega) = \sum_{i,j,I} \frac{\partial^2 a_I}{\partial x_i \partial x_j} dx_i \wedge dx_j \wedge dx_I. \quad (4.191)$$

Note that if $i = j$, then there is a repeated term in the wedge product, so

$$d(d\omega) = \sum_{i < j} \frac{\partial^2 a_I}{\partial x_i \partial x_j} dx_i \wedge dx_j \wedge dx_I \quad (4.192)$$

$$+ \sum_{i > j} \frac{\partial^2 a_I}{\partial x_i \partial x_j} dx_i \wedge dx_j \wedge dx_I. \quad (4.193)$$

Note that $dx_i \wedge dx_j = -dx_j \wedge dx_i$. We relabel the second summand to obtain

$$\begin{aligned} d(d\omega) &= \sum_{i < j} \left(\underbrace{\frac{\partial^2 a_I}{\partial x_i \partial x_j} - \frac{\partial^2 a_I}{\partial x_j \partial x_i}}_0 \right) dx_i \wedge dx_j \wedge dx_I \\ &= 0. \end{aligned} \quad (4.194)$$

□

Definition 4.42. A k -form $\omega \in \Omega^k(U)$ is *decomposable* if $\omega = \mu_1 \wedge \cdots \wedge \mu_k$, where each $\mu_i \in \Omega^1(U)$.

Theorem 4.43. If ω is decomposable, then

$$d\omega = \sum_{i=1}^k (-1)^{i-1} \mu_1 \wedge \cdots \wedge \mu_{i-1} \wedge d\mu_i \wedge \mu_{i+1} \wedge \cdots \wedge \mu_k. \quad (4.195)$$

Proof. The proof is by induction.

The case $k = 1$ is obvious. We show that if the theorem is true for $k - 1$, then the theorem is true for k .

$$\begin{aligned}
d((\mu_1 \wedge \cdots \wedge \mu_{k-1}) \wedge \mu_k) &= (d(\mu_1 \wedge \cdots \wedge \mu_{k-1})) \wedge \mu_k \\
&\quad + (-1)^{k-1}(\mu_1 \wedge \cdots \wedge \mu_{k-1}) \wedge d\mu_k \\
&= \sum_{i=1}^{k-1} (-1)^{i-1} \mu_1 \wedge \cdots \wedge d\mu_i \wedge \cdots \wedge \mu_{k-1} \wedge \mu_k \\
&\quad + (-1)^{k-1}(\mu_1 \wedge \cdots \wedge \mu_{k-1} \wedge \mu_k) \\
&= \sum_{i=1}^k (-1)^{i-1} \mu_1 \wedge \cdots \wedge d\mu_i \wedge \cdots \wedge \mu_k.
\end{aligned} \tag{4.196}$$

□

4.10 Pullback Operation on Exterior Forms

Another important operation in the theory of exterior forms is the *pullback operator*. This operation is not introduced in 18.01 or 18.02, because vector calculus is not usually taught rigorously.

Let U be open in \mathbb{R}^n and V be open in \mathbb{R}^m , and let $f : U \rightarrow V$ be a \mathcal{C}^∞ map. We can write out in components $f = (f_1, \dots, f_n)$, where each $f_i \in \mathcal{C}^\infty(U)$. Let $p \in U$ and $q = f(p)$.

The pullback of the map $df_p : T_p\mathbb{R}^m \rightarrow T_q\mathbb{R}^n$ is

$$(df_p)^* : \Lambda^k(T_q^*\mathbb{R}^n) \rightarrow \Lambda^k(T_p^*\mathbb{R}^m). \tag{4.197}$$

Suppose you have a k -form ω on V .

$$\omega \in \Omega^k(V), \tag{4.198}$$

$$\omega_q \in \Lambda^k(T_q^*\mathbb{R}^n). \tag{4.199}$$

Then

$$(df_p)^*\omega_q \in \Lambda^k(T_p^*\mathbb{R}^m). \tag{4.200}$$

Definition 4.44. $f^*\omega$ is the k -form whose value at $p \in U$ is $(df_p)^*\omega_q$.

We consider two examples. Suppose $\phi \in \Omega^0(V) = \mathcal{C}^\infty(V)$. Then $f^*\phi(p) = \phi(q)$, so $f^*\phi = \phi \circ f$, where $f : U \rightarrow V$ and $\phi : V \rightarrow \mathbb{R}$.

Again, suppose that $\phi \in \Omega^0(V) = \mathcal{C}^\infty(V)$. What is $f^*d\phi$? Let $f(p) = q$. We have the map $d\phi_q : T_q\mathbb{R}^n \rightarrow T_c\mathbb{R} = \mathbb{R}$, where $c = \phi(q)$. So,

$$\begin{aligned}
(df_p)^*(d\phi)_q &= d\phi_q \circ df_p \\
&= d(\phi \circ f)_p.
\end{aligned} \tag{4.201}$$

Therefore,

$$f^*d\phi = df^*\phi. \quad (4.202)$$

Suppose that $\mu \in \Omega^k(V)$ and $\nu \in \Omega^e(V)$. Then

$$\begin{aligned} (f^*(\mu \wedge \nu))_p &= (df_p)^*(\mu_q \wedge \nu_q) \\ &= (df_p)^*\mu_q \wedge (df_p)^*\nu_q. \end{aligned} \quad (4.203)$$

Hence,

$$f^*(\mu \wedge \nu) = f^*\mu \wedge f^*\nu. \quad (4.204)$$

We now obtain a coordinate formula for f^* .

Take $\omega \in \Omega^k(V)$. We can write $\omega = \sum a_I dx_{i_1} \wedge \cdots \wedge dx_{i_k}$, where each $a_I \in \mathcal{C}^\infty(U)$. Then

$$\begin{aligned} f^*\omega &= \sum f^*a_I f^*dx_{i_1} \wedge \cdots \wedge f^*dx_{i_k} \\ &= \sum f^*a_I df_{i_1} \wedge \cdots \wedge df_{i_k}, \end{aligned} \quad (4.205)$$

where we used the result that $f^*dx_i = dx_i \circ f = df_i$.

Note that $df_i = \sum \frac{\partial f_i}{\partial x_j} dx_j$, where $\frac{\partial f_i}{\partial x_j} \in \mathcal{C}^\infty(U)$. Also, $f^*a_I = a_I \circ f \in \mathcal{C}^\infty(U)$, which shows that

$$f^*\omega \in \Omega^k(U). \quad (4.206)$$

The following theorem states a very useful property of the pullback operator.

Theorem 4.45. *Let $\omega \in \Omega^k(V)$. Then,*

$$df^*\omega = f^*d\omega. \quad (4.207)$$

Proof. We have already checked this for $\omega = \phi \in \mathcal{C}^\infty(V)$, $k = 0$ already. We now prove the general case.

We can write $\omega = \sum a_I dx_I$. Then

$$f^*\omega = \sum f^*a_I df_{i_1} \wedge \cdots \wedge df_{i_k}. \quad (4.208)$$

So,

$$\begin{aligned} df^*\omega &= \sum df^*a_I \wedge df_{i_1} \wedge \cdots \wedge df_{i_k} \\ &\quad + \sum f^*a_I \wedge d(df_{i_1} \wedge \cdots \wedge df_{i_k}) \end{aligned} \quad (4.209)$$

Note that

$$d(df_{i_1} \wedge \cdots \wedge df_{i_k}) = \sum_{r=1}^k (-1)^{r-1} df_{i_1} \wedge \cdots \wedge d(df_{i_r}) \wedge \cdots \wedge df_{i_k}. \quad (4.210)$$

We know that $d(df_{i_r}) = 0$, so

$$\begin{aligned}df^*\omega &= \sum_I df^*a_I \wedge df_{i_1} \wedge \cdots \wedge df_{i_k} \\&= \sum_I f^*da_I \wedge f^*(dx_{i_1} \wedge \cdots \wedge dx_{i_k}) \\&= f^*\left(\sum da_I \wedge dx_I\right) \\&= f^*d\omega.\end{aligned}\tag{4.211}$$

□

Lecture 24

We review the pullback operation from last lecture. Let U be open in \mathbb{R}^m and let V be open in \mathbb{R}^n . Let $f : U \rightarrow V$ be a \mathcal{C}^∞ map, and let $f(p) = q$. From the map

$$df_p : T_p\mathbb{R}^m \rightarrow T_q\mathbb{R}^n, \quad (4.212)$$

we obtain the pullback map

$$\begin{aligned} (df_p)^* : \Lambda^k(T_q^*) &\rightarrow \Lambda^k(T_p^*) \\ \omega &\in \Omega^k(V) \rightarrow f^*\omega \in \Omega^k(U). \end{aligned} \quad (4.213)$$

We define, $f^*\omega_p = (df_p)^*\omega_q$, when $\omega_q \in \Lambda^k(T_q^*)$.

The pullback operation has some useful properties:

1. If $\omega_i \in \Omega^{k_i}(V)$, $i = 1, 2$, then

$$f^*(\omega_1 \wedge \omega_2) = f^*\omega_1 \wedge f^*\omega_2. \quad (4.214)$$

2. If $\omega \in \Omega^k(V)$, then

$$df^*\omega = f^*d\omega. \quad (4.215)$$

We prove some other useful properties of the pullback operation.

Claim. For all $\omega \in \Omega^k(W)$,

$$f^*g^*\omega = (g \circ f)^*\omega. \quad (4.216)$$

Proof. Let $f(p) = q$ and $g(q) = w$. We have the pullback maps

$$(df_p)^* : \Lambda^k(T_q^*) \rightarrow \Lambda^k(T_p^*) \quad (4.217)$$

$$(dg_q)^* : \Lambda^k(T_w^*) \rightarrow \Lambda^k(T_q^*) \quad (4.218)$$

$$(g \circ f)^* : \Lambda^k(T_w^*) \rightarrow \Lambda^k(T_p^*). \quad (4.219)$$

The chain rule says that

$$(dg \circ f)_p = (dg)_q \circ (df)_p, \quad (4.220)$$

so

$$d(g \circ f)_p^* = (df_p)^*(dg_q)^*. \quad (4.221)$$

□

Let U, V be open sets in \mathbb{R}^n , and let $f : U \rightarrow V$ be a \mathcal{C}^∞ map. We consider the pullback operation on n -forms $\omega \in \Omega^n(V)$. Let $f(0) = q$. Then

$$(dx_i)_p, \quad i = 1, \dots, n, \quad \text{is a basis of } T_p^*, \text{ and} \quad (4.222)$$

$$(dx_i)_q, \quad i = 1, \dots, n, \quad \text{is a basis of } T_q^*. \quad (4.223)$$

Using $f_i = x_i \circ f$,

$$\begin{aligned} (df_p)^*(dx_i)_q &= (df_i)_p \\ &= \sum \frac{\partial f_i}{\partial x_j}(p)(dx_j)_p. \end{aligned} \quad (4.224)$$

In the Multi-linear Algebra notes, we show that

$$(df_p)^*(dx_1)_q \wedge \cdots \wedge (dx_n)_q = \det \left[\frac{\partial f_i}{\partial x_j}(p) \right] (dx_1)_p \wedge \cdots \wedge (dx_n)_p. \quad (4.225)$$

So,

$$f^* dx_1 \wedge \cdots \wedge dx_n = \det \left[\frac{\partial f_i}{\partial x_j} \right] dx_1 \wedge \cdots \wedge dx_n. \quad (4.226)$$

Given $\omega = \phi(x)dx_1 \wedge \cdots \wedge dx_n$, where $\phi \in \mathcal{C}^\infty$,

$$f^*\omega = \phi(f(x)) \det \left[\frac{\partial f_i}{\partial x_j} \right] dx_1 \wedge \cdots \wedge dx_n. \quad (4.227)$$

5 Integration with Differential Forms

Let U be an open set in \mathbb{R}^n , and let $\omega \in \Omega^k(U)$ be a differential k -form.

Definition 5.1. The *support* of ω is

$$\text{supp } \omega = \overline{\{p \in U : \omega_p \neq 0\}}. \quad (5.1)$$

Definition 5.2. The k -form ω is *compactly supported* if $\text{supp } \omega$ is compact. We define

$$\Omega_c^k(U) = \text{the space of all compactly supported } k\text{-forms.} \quad (5.2)$$

Note that

$$\Omega_c^0(U) = \mathcal{C}_0^\infty(\mathbb{R}^n). \quad (5.3)$$

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

$$\omega = \phi(x)dx_1 \wedge \cdots \wedge dx_n, \quad (5.4)$$

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

Definition 5.3.

$$\int_U \omega \equiv \int_U \phi = \int_U \phi(x)dx_1 \cdots dx_n. \quad (5.5)$$

We are going to state and prove the change of variables theorem for integrals of differential k -forms. To do so, we first need the notions of *orientation preserving* and *orientation reversing*.

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

$$Df(p) \cong \left[\frac{\partial f_i}{\partial x_j}(p) \right]. \quad (5.6)$$

The map f is a diffeomorphism, so

$$\det \left[\frac{\partial f_i}{\partial x_j}(p) \right] \neq 0. \quad (5.7)$$

So, if U is connected, then this determinant is either positive everywhere or negative everywhere.

Definition 5.4. The map f is *orientation preserving* if $\det > 0$ everywhere. The map f is *orientation reversing* if $\det < 0$ everywhere.

The following is the change of variables theorem:

Theorem 5.5. *If $\omega \in \Omega_c^n(V)$, then*

$$\int_U f^* \omega = \int_V \omega \quad (5.8)$$

if f is orientation preserving, and

$$\int_U f^* \omega = - \int_V \omega \quad (5.9)$$

if f is orientation reversing.

In Munkres and most texts, this formula is written in slightly uglier notation. Let $\omega = \phi(x) dx_1 \wedge \cdots \wedge dx_n$, so

$$f^* \omega = \phi(f(x)) \det \left[\frac{\partial f_i}{\partial x_j} \right] dx_1 \wedge \cdots \wedge dx_n. \quad (5.10)$$

The theorem can be written as following:

Theorem 5.6. *If f is orientation preserving, then*

$$\int_V \phi = \int_U \phi \circ f \det \left[\frac{\partial f_i}{\partial x_j} \right]. \quad (5.11)$$

This is the coordinate version of the theorem.

We now prove a useful theorem found in the Supplementary Notes (and Spivak) called Sard's Theorem.

Let U be open in \mathbb{R}^n , and let $f : U \rightarrow \mathbb{R}^n$ be a $\mathcal{C}^1(U)$ map. For every $p \in U$, we have the map $Df(p) : \mathbb{R}^n \rightarrow \mathbb{R}^n$. We say that p is a *critical point of f* if $Df(p)$ is *not* bijective. Denote

$$C_f = \text{the set of all critical points of } f. \quad (5.12)$$

Sard's Theorem. *The image $f(C_f)$ is of measure zero.*

Proof. The proof is in the Supplementary Notes. □

As an example of Sard's Theorem, let $c \in \mathbb{R}^n$ and let $f : U \rightarrow \mathbb{R}^n$ be the map defined by $f(x) = c$. Note that $Df(p) = 0$ for all $p \in U$, so $C_f = U$. The set $C_f = U$ is not a set of measure zero, but $f(C_f) = \{c\}$ is a set of measure zero.

As an exercise, you should prove the following claim:

Claim. *Sard's Theorem is true for maps $f : U \rightarrow \mathbb{R}^n$, where U is an open, connected subset of \mathbb{R} .*

Proof Hint: Let $f \in \mathcal{C}^\infty(U)$ and define $g = \frac{\partial f}{\partial x}$. The map g is continuous because $f \in \mathcal{C}^1(U)$. Let $I = [a, b] \subseteq U$, and define $\ell = b - a$. The continuity of g implies that g is uniformly continuous on I . That is, for every $\epsilon > 0$, there exists a number $N > 0$ such that $|g(x) - g(y)| < \epsilon$ whenever $x, y \in I$ and $|x - y| < \ell/N$.

Now, slice I into N equal subintervals. Let $I_r, r = 1, \dots, k \leq N$ be the subintervals intersecting C_f . Prove the following lemma:

Lemma 5.7. *If $x, y \in I_r$, then $|f(x) - f(y)| < \epsilon \ell/N$.*

Proof Hint: Find $c \in I_r$ such that $f(x) - f(y) = (x - y)g(c)$. There exists $c_0 \in I_r \cap C_f$ if and only if $g(c_0) = 0$. So, we can take

$$|g(c)| = |g(c) - g(c_0)| \leq \epsilon. \quad (5.13)$$

Then $|f(x) - f(y)| \leq \epsilon \ell/N$. □

From the lemma, we can conclude that

$$f(I_r) \equiv J_r \quad (5.14)$$

is of length less than $\epsilon \ell/N$. Therefore,

$$f(C_f \cap I) \subset \bigcup_{r=1}^k J_r \quad (5.15)$$

is of length less than

$$\frac{\epsilon \ell}{N} k \leq \frac{\epsilon \ell N}{N} = \epsilon \ell. \quad (5.16)$$

Letting $\epsilon \rightarrow 0$, we find that $F(C_f \cap I)$ is of measure zero.

To conclude the proof, let $I_m, m = 1, 2, 3, \dots$, be an exhaustion of U by closed intervals $I_1 \subset I_2 \subset I_3 \subset \dots$ such that $\bigcup I_m = U$. We have shown that $f(C_f \cap I_m)$ is measure zero. So, $f(C_f) = \bigcup f(C_f \cap I_m)$ implies that $f(C_f)$ is of measure zero. \square

Lecture 25

5.1 The Poincare Lemma

Let U be an open subset of \mathbb{R}^n , and let $\omega \in \Omega^k(U)$ be a k -form. We can write $\omega = \sum a_I dx_I$, $I = (i_1, \dots, i_k)$, where each $a_I \in \mathcal{C}^\infty(U)$. Note that

$$\omega \in \Omega_c^k \iff a_I \in \mathcal{C}_0^\infty(U) \text{ for each } I. \quad (5.17)$$

We are interested in $\omega \in \Omega_c^n(U)$, which are of the form

$$\omega = f dx_1 \wedge \dots \wedge dx_n, \quad (5.18)$$

where $f \in \mathcal{C}_0^\infty(U)$. We define

$$\int_U \omega = \int_U f = \int_U f dx, \quad (5.19)$$

the Riemann integral of f over U .

Our goal over the next couple lectures is to prove the following fundamental theorem known as the Poincare Lemma.

Poincare Lemma. *Let U be a connected open subset of \mathbb{R}^n , and let $\omega \in \Omega_c^n(U)$. The following conditions are equivalent:*

1. $\int_U \omega = 0$,
2. $\omega = d\mu$, for some $\mu \in \Omega_c^{n-1}(U)$.

In today's lecture, we prove this for $U = \text{Int } Q$, where $Q = [a_1, b_1] \times \dots \times [a_n, b_n]$ is a rectangle.

Proof. First we show that (2) implies (1).

Notation.

$$dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_n \equiv dx_1 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_n. \quad (5.20)$$

Let $\mu \in \Omega_c^{n-1}(U)$. Specifically, define

$$\mu = \sum_i f_i dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_n, \quad (5.21)$$

where each $f_i \in \mathcal{C}_0^\infty(U)$. Every $\mu \in \Omega_c^{n-1}(U)$ can be written this way.

Applying d we obtain

$$d\mu = \sum_i \sum_j \frac{\partial f_i}{\partial x_j} dx_j \wedge dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_n. \quad (5.22)$$

Notice that if $i \neq j$, then the i, j th summand is zero, so

$$\begin{aligned} d\mu &= \sum_i \frac{\partial f_i}{\partial x_i} dx_i \wedge dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n \\ &= \sum (-1)^{i-1} \frac{\partial f_i}{\partial x_i} dx_1 \wedge \cdots \wedge dx_n. \end{aligned} \quad (5.23)$$

Integrate to obtain

$$\int_U d\mu = \sum (-1)^{i-1} \int_U \frac{\partial f_i}{\partial x_i}. \quad (5.24)$$

Note that

$$\int_{a_i}^{b_i} \frac{\partial f_i}{\partial x_i} dx_i = f_i(x)|_{x_i=a_i}^{x_i=b_i} = 0 - 0 = 0, \quad (5.25)$$

because f is compactly supported in U . It follows from the Fubini Theorem that

$$\int_U \frac{\partial f_i}{\partial x_i} = 0. \quad (5.26)$$

Now we prove the other direction, that (1) implies (2). Before our proof we make some remarks about functions of one variable.

Suppose $I = (a, b) \subseteq \mathbb{R}$, and let $g \in C_0^\infty(I)$ such that $\text{supp } g \subseteq [c, d]$, where $a < c < d < b$. Also assume that

$$\int_a^b g(s) ds = 0. \quad (5.27)$$

Define

$$h(x) = \int_a^x g(s) ds, \quad (5.28)$$

where $a \leq x \leq b$.

Claim. *The function h is also supported on c, d .*

Proof. If $x > d$, then we can write

$$h(x) = \int_a^b g(s) ds - \int_x^b g(s) ds, \quad (5.29)$$

where the first integral is zero by assumption, and the second integral is zero because the integrand is zero. \square

Now we begin our proof that (1) implies (2).

Let $\omega \in \Omega^n(U)$, where $U = Q$, and assume that

$$\int_U \omega = 0. \quad (5.30)$$

We will use the following inductive lemma:

Lemma 5.8. For all $0 \leq k \leq n + 1$, there exists $\mu \in \Omega_c^{n-1}(U)$ and $f \in \mathcal{C}_0^\infty(U)$ such that

$$\omega = d\mu + f dx_1 \wedge \cdots \wedge dx_n \quad (5.31)$$

and

$$\int f(x_1, \dots, x_n) dx_k \dots dx_n = 0. \quad (5.32)$$

Note that the hypothesis for $k = 0$ and $\mu = 0$ says that $\int \omega = 0$, which is our assumption (1). Also note that the hypothesis for $k = n + 1$, $f = 0$, and $\omega = d\mu$ is the same as the statement (2). So, if we can show that (the lemma is true for k) implies (the lemma is true for $k + 1$), then we will have shown that (1) implies (2) in Poincaré's Lemma. We now show this.

Assume that the lemma is true for k . That is, we have

$$\omega = d\mu + f dx_1 \wedge \cdots \wedge dx_n \quad (5.33)$$

and

$$\int f(x_1, \dots, x_n) dx_k \dots dx_n = 0, \quad (5.34)$$

where $\mu \in \Omega_c^{n-1}(U)$, and $f \in \mathcal{C}_0^\infty(\mathbb{R})$.

We can assume that μ and f are supported on $\text{Int } Q'$, where $Q' \subseteq \text{Int } Q$ and $Q' = [c_1, d_1] \times \cdots \times [c_n, d_n]$.

Define

$$g(x_1, \dots, x_k) = \int f(x_1, \dots, x_n) dx_{k+1} \dots dx_n. \quad (5.35)$$

Note that g is supported on the interior of $[c_1, d_1] \times \cdots \times [c_k, d_k]$. Also note that

$$\int_{a_k}^{b_k} g(x_1, \dots, x_{k-1}, s) ds = \int f(x_1, \dots, x_n) dx_k \dots dx_n = 0 \quad (5.36)$$

by our assumption that the lemma holds true for k .

Now, define

$$h(x_1, \dots, x_k) = \int_{a_k}^{x_k} g(x_1, \dots, x_{k-1}, s) ds. \quad (5.37)$$

From our earlier remark about functions of one variable, h is supported on $c_k \leq x_k \leq d_k$. Also, note that h is supported on $c_i \leq x_i \leq d_i$, for $1 \leq i \leq k - 1$. We conclude therefore that h is also supported on $[c_1, d_1] \times \cdots \times [c_k, d_k]$.

Both g and its "anti-derivative" are supported.

$$\frac{\partial h}{\partial x_k} = g. \quad (5.38)$$

Let $\ell = n - k$, and consider $\rho = \rho(x_{k+1}, \dots, x_n) \in \mathcal{C}_0^\infty(\mathbb{R}^\ell)$. Assume that ρ is supported on the rectangle $[c_{k+1}, d_{k+1}] \times \cdots \times [c_n, d_n]$ and that

$$\int \rho dx_{k+1} \dots dx_n = 1. \quad (5.39)$$

We can always find such a function, so we just fix one such function.

Define

$$\nu = (-1)^k h(x_1, \dots, x_k) \rho(x_{k+1}, \dots, x_n) dx_1 \wedge \dots \wedge \widehat{dx_k} \wedge \dots \wedge dx_n. \quad (5.40)$$

The form ν is supported on $Q' = [c_1, d_1] \times \dots \times [c_n, d_n]$.

Now we compute $d\nu$,

$$d\nu = (-1)^k \sum_j \frac{\partial}{\partial x_j} (h\rho) dx_j \wedge dx_1 \wedge \dots \wedge \widehat{dx_k} \wedge \dots \wedge dx_n. \quad (5.41)$$

Note that if $j \neq k$, then the summand is zero, so

$$\begin{aligned} d\nu &= (-1)^k \frac{\partial h}{\partial x_k} \rho dx_k \wedge dx_1 \wedge \dots \wedge \widehat{dx_k} \wedge \dots \wedge dx_n \\ &= (-1) \frac{\partial h}{\partial x_k} \rho dx_1 \wedge \dots \wedge dx_n \\ &= -g \rho dx_1 \wedge \dots \wedge dx_n. \end{aligned} \quad (5.42)$$

Now, define

$$\mu_{new} = \mu - \nu, \quad (5.43)$$

and

$$f_{new} = f(x_1, \dots, x_n) - g(x_1, \dots, x_k) \rho(x_{k+1}, \dots, x_n). \quad (5.44)$$

$$\begin{aligned} \omega &= d\mu_{new} + f_{new} dx_1 \wedge \dots \wedge dx_n \\ &= d\mu + (g(x_1, \dots, x_k) \rho(x_{k+1}, \dots, x_n) - f(x_1, \dots, x_k) - g\rho) dx_1 \wedge \dots \wedge dx_n \\ &= d\mu + f dx_1 \wedge \dots \wedge dx_n \\ &= \omega. \end{aligned} \quad (5.45)$$

Note that

$$\begin{aligned} \int f_{new} &= \int f_{new}(x_1, \dots, x_n) dx_{k+1} \dots dx_n \\ &= \int f(x_1, \dots, x_n) dx_{k+1} \dots dx_n \\ &\quad - g(x_1, \dots, x_k) \int \rho(x_{k+1}, \dots, x_n) dx_{k+1} \dots dx_n \\ &= g(x_1, \dots, x_k) - g(x_1, \dots, x_k) = 0, \end{aligned} \quad (5.46)$$

which implies that the lemma is true for $k + 1$. □

Remark. In the above proof, we implicitly assumed that if $f \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, then

$$g(x_1, \dots, x_k) = \int f(x_1, \dots, x_n) dx_{k+1} \dots dx_m \quad (5.47)$$

is in $\mathcal{C}_0^\infty(\mathbb{R}^k)$. We checked the support, but we did not check that g is in $\mathcal{C}^\infty(\mathbb{R}^k)$. The proof of this is in the Supplementary Notes.

Lecture 26

We continue our study of forms with compact support. Let us begin with a review. Let $U \in \mathbb{R}^n$ be open, and let

$$\omega = \sum_I f_I(x_1, \dots, x_n) dx_I, \quad (5.48)$$

where $I = (i_1, \dots, i_k)$ is strictly increasing and $dx_I = dx_{i_1} \wedge \dots \wedge dx_{i_k}$. Then

$$\omega \text{ is compactly supported} \iff \text{every } f_I \text{ is compactly supported.} \quad (5.49)$$

By definition,

$$\text{supp } f_I = \overline{\{x \in U : f_I(x) \neq 0\}}. \quad (5.50)$$

We assume that the f_I 's are \mathcal{C}^2 maps.

Notation.

$$\Omega_c^k(U) = \text{space of compactly supported differentiable } k\text{-forms on } U. \quad (5.51)$$

Now, let $\omega \in \Omega_c^n(U)$ defined by

$$\omega = f(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n, \quad (5.52)$$

where $f \in \Omega_c^0(U)$. Then

$$\int_{\mathbb{R}^n} \omega = \int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \wedge \dots \wedge dx_n. \quad (5.53)$$

Last time we proved the Poincare Lemma for open rectangles R in \mathbb{R}^n . We assumed that $\omega \in \Omega_c^n(\text{Int } R)$. That is, we assumed that $\omega \in \Omega_c^n(\mathbb{R}^n)$ such that $\text{supp } \omega \subset \text{Int } R$. We showed that for such ω the following two conditions are equivalent:

1. $\int_{\mathbb{R}^n} \omega = 0$,
2. There exists a $\mu \in \Omega_c^{n-1}(\text{Int } R)$ such that $d\mu = \omega$.

Definition 5.9. Whenever $\omega \in \Omega^k(U)$ and $\omega = d\mu$ for some $\mu \in \Omega^{k-1}(U)$, we say that ω is *exact*.

Definition 5.10. Whenever $\omega \in \Omega^k(U)$ such that $d\omega = 0$, we say that ω is *closed*.

Observe that

$$\omega \in \Omega_c^n(U) \implies d\omega = 0. \quad (5.54)$$

Now we prove the Poincare Lemma for open connected subsets of \mathbb{R}^n .

Poincare Lemma. Let U be a connected open subset of \mathbb{R}^n , and let $\omega \in \Omega_c^n(U)$. The following conditions are equivalent:

1. $\int_U \omega = 0$,
2. $\omega = d\mu$, for some $\mu \in \Omega_c^{n-1}(U)$.

Proof. We prove this more general case by reducing the proof to the case where U is a rectangle, which we proved in the previous lecture.

First we prove that (2) implies (1). We can choose a family of rectangles $\{R_i, i \in \mathbb{N}\}$ such that

$$U = \bigcup_{i \in \mathbb{N}} \text{Int } R_i \quad (5.55)$$

Since the support of μ is compact, the set $\text{supp } \mu$ is covered by finitely many of the rectangles.

We take a partition of unity $\{\phi_i, i \in \mathbb{N}\}$ subordinate to $\{R_i\}$, so that

$$\mu = \sum_{i=1}^N \underbrace{\phi_i \mu}_{\text{supported on Int } R_i} \quad (5.56)$$

Then

$$\int d\mu = \sum_i \int_{\mathbb{R}^n} d(\phi_i \mu). \quad (5.57)$$

Each term on the r.h.s is zero by the Poincare Lemma we proved last lecture.

We now prove the other direction, that (1) implies (2). It is equivalent to show that if $\omega_1, \omega_2 \in \Omega_c^n(U)$ such that

$$\int \omega_1 = \int \omega_2, \quad (5.58)$$

then $\omega_1 \sim \omega_2$, meaning that there exists a form $\mu \in \Omega_c^{n-1}(U)$ such that $\omega_1 = \omega_2 + d\mu$.

Choose a partition of unity $\{\phi_i\}$ as before. Then

$$\omega = \sum_{i=1}^M \underbrace{\phi_i \omega}_{\text{supported on Int } R_i} \quad (5.59)$$

Let

$$\int \omega = c \in \mathbb{R}, \quad (5.60)$$

and

$$\int \phi_i \omega = c_i. \quad (5.61)$$

Choose a form ω_0 such that

$$\int \omega_0 = 1 \quad (5.62)$$

and such that $\text{supp } \omega_0 \subseteq Q_0 = R_j$ for some j . Then

$$\int \underbrace{\phi_i \omega}_{\text{supported in } R_i} = \int \underbrace{c_i \omega_0}_{\text{supported in } Q_0} \quad (5.63)$$

We want to show that there exists $\mu_i \in \Omega_c^{n-1}(U)$ such that $\phi_i \omega = c_i \omega_i + d\mu_i$.

Now we use the fact that U is connected. We use the following lemma.

Lemma 5.11. *Let U be connected. Given rectangles R_i such that $\text{supp } \phi_i \omega \subset \text{Int } R_i$, and given a fixed rectangle Q_0 and any point $x \in U$, there exists a finite sequence of rectangles R_0, \dots, R_N with the following properties: $Q_0 = R_0$, $x \in \text{Int } R_N$, and $(\text{Int } R_i) \cap (\text{Int } R_{i+1})$ is non-empty.*

We omit the proof of this lemma.

Now, define $\omega_i = \phi_i \omega$, so

$$\int \omega_i = \int c_i \omega_0. \quad (5.64)$$

Note that

$$\text{supp } (c_i \omega_0) \subseteq \text{Int } (Q_0) \quad (5.65)$$

$$\text{supp } (\omega_i) \subseteq \text{Int } (R_i). \quad (5.66)$$

Choose forms ν_i such that $\text{supp } \nu_i \subseteq \text{Int } R_i \cap \text{Int } R_{i+1}$ and such that

$$\int \nu_i = 1. \quad (5.67)$$

This implies that

$$\text{supp } (\nu_i - \nu_{i+1}) \subseteq \text{Int } R_{i+1} \quad (5.68)$$

By definition,

$$\int (\nu_i - \nu_{i+1}) = 0. \quad (5.69)$$

By the Poincare Lemma we proved last time, $\nu_i \sim \nu_{i+1}$, so there exists $\mu_i \in \Omega_c^{n-1}(U)$ such that $\nu_i = \nu_{i+1} + d\mu_i$.

So,

$$c_i \omega_0 \sim c_i \nu_0 \sim c_i \nu_1 \sim \dots \sim c_i \nu_N \sim \phi_i \omega. \quad (5.70)$$

□

5.2 Proper Maps and Degree

We introduce a class of functions that remain compactly supported under the pullback operation.

Definition 5.12. Let $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^k$, and let $f : U \rightarrow V$ 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 $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^k$, and let $f : U \rightarrow V$ be a continuous map. Also let $\omega \in \Omega^k(V)$. The map

$$f^* : \Omega^k(V) \rightarrow \Omega^k(U) \quad (5.71)$$

is defined such that

$$\omega = g(y_1, \dots, y_n) dy_{i_1} \wedge \dots \wedge dy_{i_k} \rightarrow f^*\omega = g(f(x)) df_{i_1} \wedge \dots \wedge df_{i_k}. \quad (5.72)$$

So,

$$f^{-1}(\text{supp } \omega) \supseteq \text{supp } (f^*\omega). \quad (5.73)$$

If f is proper and $\omega \in \Omega_c^n(V)$, then $\text{supp } (f^*\omega)$ is compact, in which case the map f^* is actually of the form

$$f^* : \Omega_c^k(V) \rightarrow \Omega_c^k(U). \quad (5.74)$$

That is, $\omega \in \Omega_c^n(V) \rightarrow f^*\omega \in \Omega_c^n(U)$. So, it makes sense to take the integral

$$\int_U f^*\omega = (\deg f) \int_V \omega. \quad (5.75)$$

Theorem 5.13. Let U, V be connected open subsets of \mathbb{R}^n , and let $f : U \rightarrow V$ be a C^∞ map. For all $\omega \in \Omega_c^n(V)$,

$$\int_U f^*\omega = (\deg f) \int_V \omega. \quad (5.76)$$

Proof. Take $\omega_0 \in \Omega_c^n(V)$ such that

$$\int \omega_0 = 1. \quad (5.77)$$

Define

$$\deg f \equiv \int f^*\omega_0, \quad (5.78)$$

and suppose that

$$\int \omega = c. \quad (5.79)$$

Then

$$\int \omega = \int c\omega_0. \quad (5.80)$$

By the Poincare Lemma, $\omega \sim c\omega_0$. That is, there exists $\mu \in \Omega_c^{n-1}(V)$ such that $\omega = c\omega_0 + d\mu$. Then

$$\begin{aligned} f^*\omega &= f^*(c\omega_0) + f^*(d\mu) \\ &= f^*(c\omega_0) + d(f^*\mu), \end{aligned} \tag{5.81}$$

which shows that $f^*\omega \sim f^*(c\omega_0)$. Putting this altogether,

$$\begin{aligned} \int f^*\omega &= \int f^*(c\omega_0) \\ &= c \int f^*\omega_0 \\ &= c \deg f \\ &= \left(\int \omega \right) \deg f. \end{aligned} \tag{5.82}$$

□

We had $\omega = g(y_1, \dots, y_n)dy_1 \wedge \dots \wedge dy_n$, so

$$\begin{aligned} f^*\omega &= g(f(x))df_1 \wedge \dots \wedge df_m \\ &= g(f(x)) \det \left[\frac{\partial f_i}{\partial x_j} \right] dx_1 \wedge \dots \wedge dx_n, \end{aligned} \tag{5.83}$$

where we used the fact that

$$df_i = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j} dx_j \tag{5.84}$$

Restated in coordinates, the above theorem says that

$$\begin{aligned} \int_U g(f(x)) \det(Df) dx_1 \wedge \dots \wedge dx_n \\ = (\deg f) \int_V g(y_1, \dots, y_n) dy_1 \wedge \dots \wedge dy_n. \end{aligned} \tag{5.85}$$

Claim. Given proper maps $f : V \rightarrow W$ and $g : U \rightarrow V$, where U, V, W are connected open subsets of \mathbb{R}^n ,

$$\deg(fg) = (\deg g)(\deg f). \tag{5.86}$$

Proof. Note that $(f \circ g)^* = g^* \circ f^*$, so

$$\begin{aligned} \deg(f \circ g) \int_W \omega &= \int_U (f \circ g)^*\omega \\ &= \int_V g^*(f^*\omega) \\ &= (\deg g) \int_V f^*\omega \\ &= (\deg g)(\deg f) \int_W \omega. \end{aligned} \tag{5.87}$$

□

Lecture 27

We proved the following Poincare Lemma:

Poincare Lemma. *Let U be a connected open subset of \mathbb{R}^n , and let $\omega \in \Omega_c^n(U)$. The following conditions are equivalent:*

1. $\int_U \omega = 0$,
2. $\omega = d\mu$, for some $\mu \in \Omega_c^{n-1}(U)$.

We first proved this for the case $U = \text{Int } Q$, where Q was a rectangle. Then we used this result to generalize to arbitrary open connected sets. We discussed a nice application: proper maps and degree.

Let U, V be open subsets of \mathbb{R}^n , and let $f : U \rightarrow V$ be a \mathcal{C}^∞ map. The map f is *proper* if for every compact set $C \subseteq V$, the pre-image $f^{-1}(C)$ is also compact. Hence, if f is proper, then

$$f^*\Omega_c^k(V) \subseteq \Omega_c^k(U). \quad (5.88)$$

That is, if $\omega \in \Omega_c^k(V)$, then $f^*\omega \in \Omega_c^k(U)$, for all k .

When $k = n$,

$$\omega \in \Omega_c^n(V). \quad (5.89)$$

In which case, we compare

$$\int_v \omega \quad \text{and} \quad \int_U f^*\omega. \quad (5.90)$$

Using the Poincare Lemma, we obtain the following theorem.

Theorem 5.14. *There exists a constant γ_f with the property that for all $\omega \in \Omega_c^n(V)$,*

$$\int_U f^*\omega = \gamma_f \int_V \omega. \quad (5.91)$$

We call this constant the *degree* of f ,

Definition 5.15.

$$\text{deg}(f) = \gamma_f. \quad (5.92)$$

Let U, V, W be open connected subsets of \mathbb{R}^n , and let $f : U \rightarrow V$ and $g : V \rightarrow W$ be proper \mathcal{C}^∞ maps. Then the map $g \circ f : U \rightarrow W$ is proper, and

$$\text{deg}(g \circ f) = \text{deg}(f) \text{deg}(g). \quad (5.93)$$

Proof Hint: For all $\omega \in \Omega_c^n(W)$, $(g \circ f)^*\omega = f^*(g^*\omega)$. □

We give some examples of the degree of various maps. Let $f = T_a$, the transposition by a . That is, let $f(x) = x + a$. From #4 in section 4 of the Supplementary Notes, the map $T_a : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is proper. One can show that $\deg(T_a) = 1$.

As another example, let $f = A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a bijective linear map. Then

$$\deg A = \begin{cases} 1 & \text{if } \det A > 0, \\ -1 & \text{if } \det A < 0. \end{cases} \quad (5.94)$$

We now study the degree as it pertains to orientation preserving and orientation reversing maps.

Let U, V be connected open sets in \mathbb{R}^n , and let $f : U \rightarrow V$ be a diffeomorphism. Take $p \in U$. Then $Df(p) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is one-to-one and onto. The map f is *orientation preserving* if $\det Df(p) > 0$ for all $p \in U$, and the map f is *orientation reversing* if $\det Df(p) < 0$ for all $p \in U$.

Theorem 5.16. *If f is orientation preserving, then $\deg(f) = 1$; if f is orientation reversing, then $\deg(f) = -1$.*

Proof. Let $a \in U$ and $b = f(a)$. Define

$$f_{\text{old}} = f, \quad (5.95)$$

and define

$$f_{\text{new}} = T_{-b} \circ f_{\text{old}} \circ T_a, \quad (5.96)$$

where $T_{-b} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $T_a : \mathbb{R}^n \rightarrow \mathbb{R}^n$ are transpositions by $-b$ and a , respectively.

By the formula $\deg(g \circ f) = \deg(f) \deg(g)$,

$$\begin{aligned} \deg(f_{\text{new}}) &= \deg(T_{-b}) \deg(f_{\text{old}}) \deg(T_a) \\ &= \deg(f_{\text{old}}). \end{aligned} \quad (5.97)$$

By replacing f with f_{new} , we can assume that $0 \in U$ and $f(0) = 0$.

We can make yet another simplification, that $Df(0) = I$, the identity. To see this, let $Df(0) = A$, where $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Taking our new f , we redefine $f_{\text{old}} = f$, and we redefine $f_{\text{new}} = A^{-1} \circ f_{\text{old}}$. Then,

$$\deg(f_{\text{new}}) = \deg(A) \deg(f_{\text{old}}), \quad (5.98)$$

where

$$\begin{aligned} \deg A &= \deg(Df_{\text{old}}) \\ &= \begin{cases} 1 & \text{if } Df_{\text{old}} \text{ is orient. preserving,} \\ -1 & \text{if } Df_{\text{old}} \text{ is orient. reversing.} \end{cases} \end{aligned} \quad (5.99)$$

We again replace f with f_{new} . It suffices to prove the theorem for this new f . To summarize, we can assume that

$$0 \in U, \quad f(0) = 0. \quad \text{and} \quad Df(0) = I. \quad (5.100)$$

Consider $g(x) = x - f(x)$ (so $f(x) = x - g(x)$). Note that $(Dg)(0) = I - I = 0$. If we write $g = (g_1, \dots, g_n)$, then

$$\left[\frac{\partial g_i}{\partial x_j}(0) \right] = 0. \quad (5.101)$$

So, each $\frac{\partial g_i}{\partial x_j}(0) = 0$.

Lemma 5.17. *There exists $\delta > 0$ such that for all $|x| < \delta$,*

$$|g(x)| \leq \frac{|x|}{2}. \quad (5.102)$$

Proof. So far, we know that $g(0) = 0 - f(0) = 0$, and $\frac{\partial g_i}{\partial x_j}(0) = 0$. By continuity, there exists $\delta > 0$ such that

$$\left| \frac{\partial g_i}{\partial x_j}(x) \right| \leq \frac{1}{2n}, \quad (5.103)$$

for all $|x| < \delta$. Using the Mean-value Theorem, for all $|x| < \delta$,

$$\begin{aligned} g_i(x) &= g_i(x) - g_i(0) \\ &= \sum \frac{\partial g_i}{\partial x_j}(c) x_j, \end{aligned} \quad (5.104)$$

where $c = t_0 x$ for some $0 < t_0 < 1$. So,

$$\begin{aligned} |g_i(x)| &\leq \sum_{i=1}^n \frac{1}{2n} |x_i| \\ &\leq \frac{1}{2} \max\{|x_i|\} \\ &= \frac{1}{2} |x|. \end{aligned} \quad (5.105)$$

□

Define $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ as follows. Let $\rho \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, defined to have the following properties

$$\rho(x) = \begin{cases} 1 & \text{if } |x| < \delta/2, \\ 0 & \text{if } |x| > \delta, \\ 0 \leq \rho(x) \leq 1 & \text{otherwise.} \end{cases} \quad (5.106)$$

Remember that $f(x) = x - g(x)$. Define

$$\tilde{f} = \begin{cases} x - \rho(x)g(x) & \text{if } |x| < \delta, \\ x & \text{if } |x| > \delta. \end{cases} \quad (5.107)$$

Claim. The map \tilde{f} has the following properties:

1. $\tilde{f} = f(x)$ for all $|x| < \frac{\delta}{2}$,
2. $\tilde{f} = x$ for all $|x| > \delta$,
3. $|\tilde{f}(x)| \geq \frac{|x|}{2}$,
4. $|\tilde{f}(x)| \leq 2|x|$.

Proof. We only prove properties (3) and (4). First we prove property (3). We have $\tilde{f}(x) = x - \rho(x)g(x) = x$ when $|x| \geq \delta$, so $|\tilde{f}(x)| = |x|$ when $|x| \geq \delta$. For $|x| < \delta$, we have

$$\begin{aligned} |\tilde{f}(x)| &\geq |x - \rho(x)g(x)| \\ &= |x| - |\rho(x)g(x)| \\ &\geq |x| - \frac{|x|}{2} \\ &= \frac{|x|}{2}. \end{aligned} \tag{5.108}$$

We now prove property (4). We have $\tilde{f}(x) = x - \rho(x)g(x)$, so $|\tilde{f}(x)| = |x|$ for $x \geq \delta$. For $x < \delta$, we have

$$\begin{aligned} |\tilde{f}(x)| &\leq |x + \rho(x)g(x)| \\ &\leq |x| + \frac{1}{2}|x| \\ &\leq 2|x|. \end{aligned} \tag{5.109}$$

□

Let $Q_r \equiv \{x \in \mathbb{R}^n : |x| \leq r\}$. The student should check that

$$\text{Property (3)} \implies \tilde{f}^{-1}(Q_r) \subseteq Q_{2r} \tag{5.110}$$

and that

$$\text{Property (4)} \implies \tilde{f}^{-1}(\mathbb{R}^n - Q_{2r}) \subseteq \mathbb{R}^n - Q_r \tag{5.111}$$

Notice that $\tilde{f}^{-1}(Q_r) \subseteq Q_{2r} \implies \tilde{f}$ is proper.

Now we turn back to the map f . Remember that $f : U \rightarrow V$ is a diffeomorphism and that $f(0) = 0$. So, the set $f(\text{Int } Q_{\delta/2})$ is an open neighborhood of 0 in \mathbb{R}^n . Take

$$\omega \in \Omega_c^n(f(\text{Int } Q_{\delta/2}) \cap \text{Int } Q_{\delta/4}) \tag{5.112}$$

such that

$$\int_{\mathbb{R}^n} \omega = 1. \tag{5.113}$$

Then,

$$f^*\omega \in \Omega_c^n(Q_{\delta/2}) \quad (5.114)$$

and

$$\tilde{f}^*\omega \in \Omega_c^n(Q_{\delta/2}), \quad (5.115)$$

by Equation 5.110. This shows that $f^*\omega = \tilde{f}^*\omega$. Hence,

$$\begin{aligned} \int_U f^*\omega &= \int_U \tilde{f}^*\omega = \deg(f) \int_V \omega \\ &= \deg(\tilde{f}) \int_V \omega, \end{aligned} \quad (5.116)$$

where

$$\int_V \omega = 1. \quad (5.117)$$

Therefore,

$$\deg(f) = \deg(\tilde{f}). \quad (5.118)$$

Now, let us use Equation 5.111. Choose $\omega \in \Omega_c^n(\mathbb{R}^n - Q_{2\delta})$. So,

$$f^*\omega \in \Omega_c^n(\mathbb{R}^n - Q_\delta). \quad (5.119)$$

Again we take

$$\int_{\mathbb{R}^n} \omega = 1. \quad (5.120)$$

By property (2), $\tilde{f} = I$ on $\mathbb{R}^n - Q_\delta$, so

$$\tilde{f}^*\omega = \omega. \quad (5.121)$$

Integrating,

$$\int_{\mathbb{R}^n} \tilde{f}^*\omega = \deg(\tilde{f}) \int_{\mathbb{R}^n} \omega = \int_{\mathbb{R}^n} \omega. \quad (5.122)$$

Therefore,

$$\deg(f) = \deg(\tilde{f}) = 1. \quad (5.123)$$

□

Lecture 28

Let U, V be connected open sets of \mathbb{R}^n , and let $f : U \rightarrow V$ be a diffeomorphism. Then

$$\deg(f) = \begin{cases} +1 & \text{if } f \text{ is orient. preserving,} \\ -1 & \text{if } f \text{ is orient. reversing.} \end{cases} \quad (5.124)$$

We showed that given any $\omega \in \Omega_c^n(V)$,

$$\int_U f^* \omega = \pm \int_V \omega. \quad (5.125)$$

Let $\omega = \phi(x) dx_1 \wedge \cdots \wedge dx_n$, where $\phi \in \mathcal{C}_0^\infty(V)$. Then

$$f^* \omega = \phi(f(x)) \det \left[\frac{\partial f_i}{\partial x_j}(x) \right] dx_1 \wedge \cdots \wedge dx_n, \quad (5.126)$$

so,

$$\int_U \phi(f(x)) \det \left[\frac{\partial f_i}{\partial x_j} \right] dx = \pm \int_V \phi(x) dx. \quad (5.127)$$

Notice that

$$f \text{ is orientation preserving} \iff \det \left[\frac{\partial f_i}{\partial x_j}(x) \right] > 0, \quad (5.128)$$

$$f \text{ is orientation reversing} \iff \det \left[\frac{\partial f_i}{\partial x_j}(x) \right] < 0. \quad (5.129)$$

So, in general,

$$\int_U \phi(f(x)) \left| \det \left[\frac{\partial f_i}{\partial x_j}(x) \right] \right| dx. \quad (5.130)$$

As usual, we assumed that $f \in \mathcal{C}^\infty$.

Remark. The above is true for $\phi \in \mathcal{C}_0^1$, a compactly supported continuous function. The proof of this is in section 5 of the Supplementary Notes. The theorem is true even if only $f \in \mathcal{C}^1$ (the notes prove it for $f \in \mathcal{C}^2$).

Today we show how to compute the degree in general.

Let U, V be connected open sets in \mathbb{R}^n , and let $f : U \rightarrow V$ be a proper \mathcal{C}^∞ map.

Claim. *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 $B \subseteq V_0$ such that $f^{-1}(V_0) \subseteq U_0$.*

Proof. Let $C \subseteq V$ be a compact set with $B \subseteq \text{Int } C$, and let $W = f^{-1}(C) - U_0$. The set W is compact, so the set $f(W)$ is also compact. Moreover, $f(W) \cap B = \emptyset$ since $f^{-1}(B) \subseteq U_0$.

Now, let $V_0 = \text{Int } C - f(W)$. This set is open, and

$$\begin{aligned} f^{-1}(V_0) &\subseteq f^{-1}(\text{Int } C) - W \\ &\subseteq U_0. \end{aligned} \tag{5.131}$$

□

Claim. *If $X \subseteq U$ is closed, then $f(X)$ is closed in V .*

Proof. Take any point $p \in V - f(X)$. Then $f^{-1}(p) \in U - X$. Apply the previous result with $B = \{p\}$, $A = f^{-1}(p)$, and $U_0 = U - X$. There exists an open set $V_0 \ni p$ such that $f^{-1}(V_0) \subseteq U - X$. The set $V_0 \cap f(X) = \emptyset$, so $V - f(X)$ is open in V . □

We now remind you of Sard's Theorem. Let $f : U \rightarrow V$ be a proper \mathcal{C}^∞ map. We define the critical set

$$C_f = \{p \in U : Df(p) \text{ is not bijective}\}. \tag{5.132}$$

The set C_f is closed. The set $f(C_f)$ in V is a set of measure zero. The set $f(C_f)$ is closed as well, since f is proper.

Definition 5.18. A point $q \in V$ is a *regular value* of f if $q \in V - f(C_f)$.

Sard's Theorem basically says that there are "lots" of regular values.

Lemma 5.19. *If q is a regular value, then $f^{-1}(q)$ is a finite set.*

Proof. First, $p \in f^{-1}(q) \implies p \notin C_f$. So, $Df(p) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is bijective. By the IFT, the map f is a diffeomorphism of a neighborhood U_p of $p \in U$ onto a neighborhood of q . In particular, since f is one-to-one and onto,

$$U_p \cap f^{-1}(q) = \{p\}. \tag{5.133}$$

Consider the collection $\{U_p : p \in f^{-1}(q)\}$, which is an open cover of $f^{-1}(q)$. The H-B Theorem tells us that there exists a finite subcover $\{U_{p_i}, i = 1, \dots, N\}$. Hence,

$$f^{-1}(q) = \{p_1, \dots, p_N\}. \tag{5.134}$$

□

Theorem 5.20. *The degree of f is*

$$\deg(f) = \sum_{i=1}^N \sigma_{p_i}, \tag{5.135}$$

where

$$\sigma_{p_i} = \begin{cases} +1 & \text{if } Df(p_i) \text{ is orient. preserving,} \\ -1 & \text{if } Df(p_i) \text{ is orient. reversing.} \end{cases} \quad (5.136)$$

So, to calculate the degree, you just pick any regular value q and “count” the number of points in the pre-image of q , keeping track of the value of σ_{p_i} .

Proof. For each $p_i \in f^{-1}(q)$, let U_{p_i} be an open neighborhood of p_i such that f maps U_{p_i} diffeomorphically onto a neighborhood of q . We can assume that the U_{p_i} ’s do not intersect.

Now, choose a neighborhood V_0 of q such that

$$f^{-1}(V_0) \subseteq \bigcup U_{p_i}. \quad (5.137)$$

Next, replace each U_{p_i} by $U_{p_i} \cap f^{-1}(V_0)$. So, we can assume the following:

1. f is a diffeomorphism of U_{p_i} onto V_0 ,
2. $f^{-1}(V_0) = \bigcup U_{p_i}$,
3. The U_{p_i} ’s don’t intersect.

Choose $\omega \in \Omega_c^n(V_0)$ such that

$$\int_V \omega = 1. \quad (5.138)$$

Then,

$$\begin{aligned} \int_U f^* \omega &= \sum_i \int_{U_{p_i}} f^* \omega \\ &= \sum_i \sigma_{p_i} \int_{V_0} \omega \\ &= \sum_i \sigma_{p_i}. \end{aligned} \quad (5.139)$$

But,

$$\int_U f^* \omega = (\deg f) \int_U \omega = \deg f, \quad (5.140)$$

so

$$\sum \sigma_{p_i} = \deg f. \quad (5.141)$$

□

This is a very nice theorem that is not often discussed in textbooks.

The following is a useful application of this theorem. Suppose $f^{-1}(q)$ is empty, so $q \notin f(U)$. Then $q \notin f(C_f)$, so q is a regular value. Therefore,

$$\deg(f) = 0. \quad (5.142)$$

This implies the following useful theorem.

Theorem 5.21. *If $\deg(f) \neq 0$, then $f : U \rightarrow V$ is onto.*

This theorem can let us know if a system of non-linear equations has a solution, simply by calculating the degree. The way to think about this is as follows. Let $f = (f_1, \dots, f_n)$ and let $q = (c_1, \dots, c_n) \in V$. If $q \in f(U)$ then there exists a solution $x \in U$ to the system of non-linear equations

$$f_i(x) = c_i, \quad i = 1, \dots, n. \tag{5.143}$$

Lecture 29

We have been studying the important invariant called the degree of f . Today we show that the degree is a “topological invariant.”

5.3 Topological Invariance of Degree

Recall that given a subset A of \mathbb{R}^m and a function $F : A \rightarrow \mathbb{R}^\ell$, we say that F is \mathcal{C}^∞ if it extends to a \mathcal{C}^∞ map on a neighborhood of A .

Let U be open in \mathbb{R}^n , let V be open in \mathbb{R}^k , and let $A = U \times [0, 1]$.

Definition 5.22. Let $f_0, f_1 : U \rightarrow V$ be \mathcal{C}^∞ maps. The maps f_0 and f_1 are *homotopic* if there is a \mathcal{C}^∞ map $F : U \times [0, 1] \rightarrow V$ such that $F(p, 0) = f_0(p)$ and $F(p, 1) = f_1(p)$ for all $p \in U$.

Let $f_t : U \rightarrow V$ be the map defined by

$$f_t(p) = F(p, t). \tag{5.144}$$

Note that $F \in \mathcal{C}^\infty \implies f_t \in \mathcal{C}^\infty$. So, $f_t : U \rightarrow V$, where $0 \leq t \leq 1$, gives a family of maps parameterized by t . The family of maps f_t is called a \mathcal{C}^∞ *deformation* of f_0 into f_1 .

Definition 5.23. The map F is a *proper homotopy* if for all compact sets $A \subseteq V$, the pre-image $F^{-1}(A)$ is compact.

Denote by π the map $\pi : U \times [0, 1] \rightarrow U$ that sends $(p, t) \rightarrow p$. Let $A \subseteq V$ be compact. Then $B = \pi(F^{-1}(A))$ is compact, and for all t , $f_t^{-1}(A) \subseteq B$. As a consequence, each f_t is proper.

We concentrate on the case where U, V are open connected subsets of \mathbb{R}^n and $f_0, f_1 : U \rightarrow V$ are proper \mathcal{C}^∞ maps. We now prove that the degree is a topological invariant.

Theorem 5.24. *If f_0 and f_1 are homotopic by a proper homotopy, then*

$$\deg(f_0) = \deg(f_1). \tag{5.145}$$

Proof. Let $\omega \in \Omega_c^n(V)$ and let $\text{supp } \omega = A$. Let $F : U \times I \rightarrow V$ be a proper homotopy between f_0 and f_1 . Take $B = \pi(F^{-1}(A))$, which is compact. For all $t \in [0, 1]$, $f_t^{-1}(A) \subseteq B$.

Let us compute $f_t^* \omega$. We can write $\omega = \phi(x) dx_1 \wedge \cdots \wedge dx_n$, where $\text{supp } \phi \subseteq A$. So,

$$f_t^* \omega = \phi(F(x, t)) \det \left[\frac{\partial F_i}{\partial x_j}(x, t) \right] dx_1 \wedge \cdots \wedge dx_n, \tag{5.146}$$

and

$$\begin{aligned}\int_U f_t^* \omega &= \deg(f_t) \int_V \omega. \\ &= \int_U \phi(F(x, t)) \det \left[\frac{\partial F_i}{\partial x_j}(x, t) \right] dx_1 \dots dx_n.\end{aligned}\tag{5.147}$$

Notice that the integrand is supported in the compact set B for all t , and it is \mathcal{C}^∞ as a function of x and t . By Exercise #2 in section 2 of the Supplementary Notes, this implies that the integral is \mathcal{C}^∞ in t . From Equation 5.147, we can conclude that $\deg(f_t)$ is a \mathcal{C}^∞ function of t .

Now here is the trick. Last lecture we showed that $\deg(f_t)$ is an integer. Since $\deg(f_t)$ is continuous, it must be a constant $\deg(f_t) = \text{constant}$. \square

We consider a simple application of the above theorem. Let $U = V = \mathbb{R}^2$, and think of $\mathbb{R}^2 = \mathbb{C}$. We make the following associations:

$$i^2 = -1 \tag{5.148}$$

$$z = x + iy \tag{5.149}$$

$$\bar{z} = x - iy \tag{5.150}$$

$$z\bar{z} = |z|^2 = x^2 + y^2 \tag{5.151}$$

$$dz = dx + idy \tag{5.152}$$

$$d\bar{z} = dx - idy \tag{5.153}$$

$$dz \wedge d\bar{z} = -2idx \wedge dy \tag{5.154}$$

$$dx \wedge dy = \frac{1}{2}idz \wedge d\bar{z}. \tag{5.155}$$

Consider a map $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, thinking of $\mathbb{R}^2 = \mathbb{C}$, defined by

$$f(z) = z^n + \sum_{i=0}^{n-1} c_i z^i, \quad c_i \in \mathbb{C}. \tag{5.156}$$

Claim. *The map f is proper.*

Proof. Let $C = \sum |c_i|$. For $|z| > 1$,

$$\left| \sum_{i=0}^{n-1} c_i z^i \right| \leq C|z|^{n-1}. \tag{5.157}$$

So,

$$\begin{aligned}|f(z)| &\geq |z|^n - \left| \sum_{i=0}^{n-1} c_i z^i \right| \\ &= |z|^n - C|z|^{n-1} \\ &= |z|^n \left(1 - \frac{C}{|z|} \right).\end{aligned}\tag{5.158}$$

For $|z| > 2C$,

$$|f(z)| \geq \frac{|z|^n}{2}. \quad (5.159)$$

So, if $R > 1$ and $R > 2C$, then $f^{-1}(B_R) \subseteq B_{R_1}$, where $R_1^n/2 \leq R$ (and where B_r denotes the ball of radius r). So f is proper. \square

Now, let us define a homotopy $F : \mathbb{C} \times [0, 1] \rightarrow \mathbb{C}$ by

$$F(z, t) = z^n + t \sum_{i=0}^{n-1} c_i z^i. \quad (5.160)$$

We claim that $F^{-1}(B_R) \subseteq B_{R_1} \times [0, 1]$, by exactly the same argument as above. So F is proper.

Notice that

$$F(z, 1) = f_1(z) = f(z), \quad (5.161)$$

$$F(z, 0) = f_0(z) = z^n. \quad (5.162)$$

So, by the above theorem, $\deg(f) = \deg(f_0)$.

Let us compute $\deg(f_0)$ by brute force. We have $f_0(z) = z^n$, so

$$f_0^* dz = dz^n = n z^{n-1} dz, \quad (5.163)$$

$$f_0^* d\bar{z} = d\bar{z}^n = n \bar{z}^{n-1} d\bar{z}. \quad (5.164)$$

Using the associations defined above,

$$\begin{aligned} f_0^*(dx \wedge dy) &= \frac{i}{2} f_0^*(dz \wedge d\bar{z}) \\ &= \frac{i}{2} f_0^* dz \wedge f_0^* d\bar{z} \\ &= \frac{i}{2} n^2 |z|^{2(n-1)} dz \wedge d\bar{z} \\ &\quad + n^2 |z|^{2n-2} dx \wedge dy. \end{aligned} \quad (5.165)$$

Let $\phi \in \mathcal{C}_0^\infty(\mathbb{R})$ such that

$$\int_0^\infty \phi(s) ds = 1. \quad (5.166)$$

Let $\omega = \phi(|z|^2) dx \wedge dy$. We calculate $\int_{\mathbb{R}^2} \omega$. Let us use polar coordinates, where

$$r = \sqrt{x^2 + y^2} = |z|.$$

$$\begin{aligned} \int_{\mathbb{R}^2} \omega &= \int_{\mathbb{R}^2} \phi(|z|^2) dx dy \\ &= \int_{\mathbb{R}^2} \phi(r^2) r dr d\theta \\ &= 2\pi \int_0^\infty \phi(r^2) r dr \\ &\quad + 2\pi \int_0^\infty \phi(s) \frac{ds}{2} \\ &= \pi. \end{aligned} \tag{5.167}$$

Now we calculate $\int f_0^* \omega$. First, we note that

$$f_0^* \omega = \phi(|z|^{2n}) n^2 |z|^{2n-2} dx \wedge dy. \tag{5.168}$$

So,

$$\begin{aligned} \int f_0^* \omega &= n^2 \int_0^\infty \phi(r^{2n}) r^{2n-2} r dr d\theta \\ &= n^2 (2\pi) \int_0^\infty \phi(r^{2n}) r^{2n-1} dr \\ &= n^2 (2\pi) \int_0^\infty \phi(s) \frac{ds}{2n} \\ &= n\pi. \end{aligned} \tag{5.169}$$

To summarize, we have calculated that

$$\int_{\mathbb{R}^2} \omega = \pi \quad \text{and} \quad \int_{\mathbb{R}^2} f_0^* \omega = n\pi. \tag{5.170}$$

Therefore,

$$\deg(f_0) = \deg(f) = n. \tag{5.171}$$

A better way to do the above calculation is in the homework: problem #6 of section 6 of the Supplementary Notes.

Last lecture we showed that if $\deg(f) \neq 0$, then the map f is onto. Applying this to the above example, we find that the algebraic equation

$$z^n + \sum_{i=0}^{n-1} c_i z^i = 0 \tag{5.172}$$

has a solution. This is known as the Fundamental Theorem of Algebra.

Lecture 30

6 Manifolds

6.1 Canonical Submersion and Canonical Immersion Theorems

As part of today's homework, you are to prove the canonical submersion and immersion theorems for linear maps. We begin today's lecture by stating these two theorems.

Let $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear map, and let $[a_{ij}]$ be its associated matrix. We have the transpose map $A^t : \mathbb{R}^m \rightarrow \mathbb{R}^n$ with the associated matrix $[a_{ji}]$.

Definition 6.1. Let $k < n$. Define the *canonical submersion* map π and the *canonical immersion* map ι as follows:

Canonical submersion:

$$\pi : \mathbb{R}^n \rightarrow \mathbb{R}^k, \quad (x_1, \dots, x_n) \rightarrow (x_1, \dots, x_k). \quad (6.1)$$

Canonical immersion:

$$\iota : \mathbb{R}^k \rightarrow \mathbb{R}^n, \quad (x_1, \dots, x_k) \rightarrow (x_1, \dots, x_k, 0, \dots, 0). \quad (6.2)$$

Canonical Submersion Theorem. Let $A : \mathbb{R}^n \rightarrow \mathbb{R}^k$ be a linear map, and suppose that A is onto. Then there exists a bijective linear map $B : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $A \circ B = \pi$.

Proof Hint: Show that there exists a basis v_1, \dots, v_n of \mathbb{R}^n such that $Av_i = e_i$, $i = 1, \dots, k$, (the standard basis of \mathbb{R}^k) and $Av_i = 0$ for all $i > k$. Then let $B : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the linear map $Be_i = v_i$, $i = 1, \dots, n$, where e_1, \dots, e_n is the standard basis of \mathbb{R}^n . \square

Canonical Immersion Theorem. As before, let $k < n$. Let $A : \mathbb{R}^k \rightarrow \mathbb{R}^n$ be a one-to-one linear map. Then there exists a bijective linear map $B : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $B \circ A = \iota$.

Proof Hint: Note that $B \circ A = \iota \iff A^t B^t = \pi$. Use the Canonical Submersion Theorem. \square

Now we prove non-linear versions of these two theorems.

Let U be an open set in \mathbb{R}^n , and let $f : U \rightarrow \mathbb{R}^k$ be a \mathcal{C}^∞ map. Let $p \in U$.

Definition 6.2. The map f is a *submersion at p* if $Df(p) : \mathbb{R}^n \rightarrow \mathbb{R}^k$ is onto.

Canonical Submersion Theorem. *Assume that f is a submersion at p and that $f(p) = 0$. Then there exists a neighborhood U_0 of p in U , a neighborhood V of 0 in \mathbb{R}^n , and a diffeomorphism $g : V \rightarrow U_0$ such that $f \circ g = \pi$.*

Proof. Let $T_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the translation defined by $x \rightarrow x + p$. Replacing f by $f \circ T_p$ we can assume that $p = 0$ and $f(0) = 0$.

Let $A = (Df)(0)$, where $A : \mathbb{R}^n \rightarrow \mathbb{R}^k$ is onto by the assumption that f is a submersion. So, there exists a bijective linear map $B : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $A \circ B = \pi$. Replacing f by $f \circ B$ we can assume that $Df(0) = \pi$.

Define a map $h : U \rightarrow \mathbb{R}^n$ by

$$h(x_1, \dots, x_n) = (f(x_1, \dots, x_k); x_{k+1}, \dots, x_n). \quad (6.3)$$

Note that (1) $Dh(0) = I$; and (2) $\pi h = f$. By (1), the function h maps a neighborhood U_0 of 0 in U diffeomorphically onto a neighborhood V of 0 in \mathbb{R}^n . By (2), we have $\pi = f \circ h^{-1}$. Take $g = h^{-1}$. \square

There is a companion theorem having to do with immersions.

Definition 6.3. Let U be an open subset of \mathbb{R}^k , and let $f : U \rightarrow \mathbb{R}^n$ be a C^∞ map. Let $p \in U$. The map f is an *immersion at p* if $(Df)(p) : \mathbb{R}^k \rightarrow \mathbb{R}^n$ is injective (one-to-one).

Canonical Immersion Theorem. *Let U be a neighborhood of 0 in \mathbb{R}^k , and let $f : U \rightarrow \mathbb{R}^n$ be a C^∞ map. Assume that f is an immersion at 0 . Then there exists a neighborhood V of $f(0) = p$ in \mathbb{R}^n , a neighborhood W of 0 in \mathbb{R}^k , and a diffeomorphism $g : V \rightarrow W$ such that $\iota^{-1}(W) \subseteq U$ and $g \circ f = \iota$.*

Proof. Replacing f by $T_p \circ f$, we can assume that $f(0) = 0$. Let $A = Df(0)$, so $A : \mathbb{R}^k \rightarrow \mathbb{R}^n$ is injective. There exists a linear map $B : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $BA = \iota$. Replacing f by $B \circ f$, we can assume that $Df(0) = \iota$.

Let $\ell = n - k$. Since $U \subseteq \mathbb{R}^k$, we get $U \times \mathbb{R}^\ell \subseteq \mathbb{R}^k \times \mathbb{R}^\ell = \mathbb{R}^n$. Define a map $h : U \times \mathbb{R}^\ell \rightarrow \mathbb{R}^n$ by

$$h(x_1, \dots, x_n) = f(x_1, \dots, x_k) + (0, \dots, 0, x_{k+1}, \dots, x_n). \quad (6.4)$$

One can check that (1) $Dh(0) = I$; and (2) $h \circ \iota = f$.

By (1), the function h maps a neighborhood W of 0 in $U \times \mathbb{R}^\ell$ diffeomorphically onto a neighborhood V of 0 in \mathbb{R}^n . Moreover, $W \subseteq U \times \mathbb{R}^\ell$, so $\iota^{-1}(W) \subseteq U$.

By (2), we obtain the canonical immersion map $\iota = h^{-1} \circ f$. Take $g = h^{-1}$. \square

6.2 Definition of Manifold

Now we move on to the study of manifolds.

Let X be a subset of \mathbb{R}^n , let Y be a subset of \mathbb{R}^m , and let $f : X \rightarrow Y$ be a continuous map. We define that the map f is a \mathcal{C}^∞ map if for every point $p \in X$ there exists a neighborhood U_p of p in \mathbb{R}^n and a \mathcal{C}^∞ map $g_p : U_p \rightarrow \mathbb{R}^m$ such that $g_p|_{X \cap U_p} = f$.

We showed in the homework that if $f : X \rightarrow Y$ is a \mathcal{C}^∞ map, then there exists a neighborhood U of X in \mathbb{R}^n and a \mathcal{C}^∞ map $g : U \rightarrow \mathbb{R}^m$ extending f .

Definition 6.4. A map $f : X \rightarrow Y$ is a *diffeomorphism* if it is one-to-one, onto, a \mathcal{C}^∞ map, and $f^{-1} : Y \rightarrow X$ is \mathcal{C}^∞ .

Let X be a subset of \mathbb{R}^N .

Definition 6.5. The set X is an *n-dimensional manifold* if for every point $p \in X$ there exists a neighborhood V of p in \mathbb{R}^N , an open set U in \mathbb{R}^m , and a diffeomorphism $f : U \rightarrow V \cap X$. The collection (f, U, X) is called a *parameterization of X at p* .

This definition does not illustrate how manifolds come up in nature. Usually manifolds come up in the following scenario.

Let W be open in \mathbb{R}^N , and let $f_i : W \rightarrow \mathbb{R}$, $i = 1, \dots, \ell$ be \mathcal{C}^∞ functions. Suppose you want to study the solution space of

$$f_i(x_1, \dots, x_N) = 0, \quad i = 1, \dots, \ell. \quad (6.5)$$

Then you consider the mapping $f : W \rightarrow \mathbb{R}^\ell$ defined by

$$f(x) = (f_1(x), \dots, f_\ell(x)). \quad (6.6)$$

Claim. *If for every $p \in W$ the map f is a submersion of p , then Equation 6.6 defines a k -dimensional manifold, where $k = N - \ell$.*

Lecture 31

6.3 Examples of Manifolds

We begin with a review of the definition of a manifold.

Let X be a subset of \mathbb{R}^n , let Y be a subset of \mathbb{R}^m , and let $f : X \rightarrow Y$ be a continuous map.

Definition 6.6. The map f is \mathcal{C}^∞ if for every $p \in X$, there exists a neighborhood U_p of p in \mathbb{R}^n and a \mathcal{C}^∞ map $g_p : U_p \rightarrow \mathbb{R}^m$ such that $g_p = f$ on $U_p \cap X$.

Claim. If $f : X \rightarrow Y$ is continuous, then there exists a neighborhood U of X in \mathbb{R}^n and a \mathcal{C}^∞ map $g : U \rightarrow \mathbb{R}^m$ such that $g = f$ on $U \cap X$.

Definition 6.7. The map $f : X \rightarrow Y$ is a *diffeomorphism* if it is one-to-one, onto, and both f and f^{-1} are \mathcal{C}^∞ maps.

We define the notion of a manifold.

Definition 6.8. A subset X of \mathbb{R}^N is an *n-dimensional manifold* if for every $p \in X$, there exists a neighborhood V of p in \mathbb{R}^N , an open set U in \mathbb{R}^n , and a diffeomorphism $\phi : U \rightarrow X \cap V$.

Intuitively, the set X is an n -dimensional manifold if locally near every point $p \in X$, the set X “looks like an open subset of \mathbb{R}^n .”

Manifolds come up in practical applications as follows:

Let U be an open subset of \mathbb{R}^N , let $k < N$, and let $f : \mathbb{R}^N \rightarrow \mathbb{R}^k$ be a \mathcal{C}^∞ map. Suppose that 0 is a regular value of f , that is, $f^{-1}(0) \cap C_f = \emptyset$.

Theorem 6.9. The set $X = f^{-1}(0)$ is an n -dimensional manifold, where $n = N - k$.

Proof. If $p \in f^{-1}(0)$, then $p \notin C_f$. So the map $Df(p) : \mathbb{R}^N \rightarrow \mathbb{R}^k$ is onto. The map f is a submersion at p .

By the canonical submersion theorem, there exists a neighborhood V of 0 in \mathbb{R}^k , a neighborhood U_0 of p in U , and a diffeomorphism $g : V \rightarrow U$ such that

$$f \circ g = \pi. \tag{6.7}$$

Recall that $\mathbb{R}^N = \mathbb{R}^k \times \mathbb{R}^n$ and $\pi : \mathbb{R}^N \rightarrow \mathbb{R}^k$ is the map that sends

$$(x, y) \in \mathbb{R}^k \times \mathbb{R}^n \rightarrow \mathbb{R}^k. \tag{6.8}$$

Hence, $\pi^{-1}(0) = \{0\} \times \mathbb{R}^n = \mathbb{R}^n$. By Equation 6.7, the function g maps $V \cap \pi^{-1}(0)$ diffeomorphically onto $U_0 \cap f^{-1}(0)$. But $V \cap \pi^{-1}(0)$ is a neighborhood of 0 in \mathbb{R}^k and $U_0 \cap f^{-1}(0)$ is a neighborhood of p in X . \square

We give three examples of applications of the preceding theorem.

1. We consider the n -sphere S^n . Define a map

$$f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}, \quad f(x) = x_1^2 + \dots + x_{n+1}^2 - 1. \quad (6.9)$$

The derivative is $(Df)(x) = 2[x_1, \dots, x_{n+1}]$, so $C_f = \{0\}$. If $a \in f^{-1}(0)$, then $\sum a_i^2 = 1$, so $a \notin C_f$. Thus, the set $f^{-1}(0) = S^n$ is an n -dimensional manifold.

2. Let $g : \mathbb{R}^n \rightarrow \mathbb{R}^k$ be a C^∞ map. Define

$$X = \text{graph } g = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^k : y = g(x)\}. \quad (6.10)$$

Note that $X \subseteq \mathbb{R}^n \times \mathbb{R}^k = \mathbb{R}^{n+k}$.

Claim. *The set X is an n -dimensional manifold.*

Proof. Define a map $f : \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^k$ by

$$f(x, y) = y - g(x). \quad (6.11)$$

Note that $Df(x, y) = [-Dg(x), I_k]$. This is always of rank k , so $C_f = \emptyset$. Hence, the graph g is an n -dimensional manifold. \square

3. The following example comes from Munkres section 24, exercise #6. Let

$$\mathcal{M}_n = \text{the set of all } n \times n \text{ matrices}, \quad (6.12)$$

so

$$\mathcal{M}_n \cong \mathbb{R}^{n^2}. \quad (6.13)$$

With any element $[a_{ij}]$ in \mathcal{M}_n we associate a vector

$$(a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots). \quad (6.14)$$

Now, let

$$\mathcal{S}_n = \{A \in \mathcal{M}_n : A = A^t\}, \quad (6.15)$$

so

$$\mathcal{S}_n \cong \mathbb{R}^{\frac{n(n+1)}{2}}. \quad (6.16)$$

With any element $[a_{ij}]$ in \mathcal{S}_n we associate a vector

$$(a_{11}, \dots, a_{1n}, a_{22}, a_{23}, \dots, a_{2n}, a_{33}, a_{34}, \dots). \quad (6.17)$$

The above association avoids the “redundancies” $a_{12} = a_{21}, a_{31} = a_{13}, a_{32} = a_{23}$, etc.

Define

$$O(n) = \{A \in \mathcal{M}_n : A^t A = I\}, \quad (6.18)$$

which is the set of orthogonal $n \times n$ matrices.

As an exercise, the student should prove the following claim.

Claim. The set $O(n) \subseteq \mathcal{M}_n$ is an $\frac{n(n-1)}{2}$ -dimensional manifold.

Proof Hint: First hint: Let $f : \mathcal{M}_n \rightarrow \mathcal{S}_n$ be the map defined by

$$f(A) = A^t A - I, \quad (6.19)$$

so $O(n) = f^{-1}(0)$. Show that $f^{-1}(0) \cap C_f = \emptyset$. The main idea is to show that if $A \notin f^{-1}(0)$, then the map $Df(A) : \mathcal{M}_n \rightarrow \mathcal{S}_n$ is onto.

Second hint: Note that $Df(A)$ is the map that sends $B \in \mathcal{M}_n$ to $A^t B + B^t A$. \square

Manifolds are often defined by systems of non-linear equations:

Let $f : \mathbb{R}^N \rightarrow \mathbb{R}^k$ be a continuous map, and suppose that $C_f \cap f^{-1}(0) = \emptyset$. Then $X = f^{-1}(0)$ is an n -dimensional manifold. Suppose that $f = (f_1, \dots, f_k)$. Then X is defined by the system of equations

$$f_i(x_1, \dots, x_N) = 0, \quad i = 1, \dots, k. \quad (6.20)$$

This system of equations is called *non-degenerate*, since for every $x \in X$ the matrix

$$\left[\frac{\partial f_i}{\partial x_j}(x) \right] \quad (6.21)$$

is of rank k .

Claim. Every n -dimensional manifold $X \subseteq \mathbb{R}^N$ can be described locally by a system of k non-degenerate equations of the type above.

Proof Idea: Let $X \subseteq \mathbb{R}^N$ be an n -dimensional manifold. Let $p \in X$, let U be an open subset of \mathbb{R}^n , and let V be an open neighborhood of p in \mathbb{R}^N . Let $\phi : U \rightarrow V \cap X$ be a diffeomorphism. Modifying ϕ by a translation if necessary we can assume that $0 \in U$ and $\phi(0) = p$. We can think of ϕ as a map $\phi : U \rightarrow \mathbb{R}^N$ mapping U into X .

Claim. The linear map $(D\phi)(0) : \mathbb{R}^n \rightarrow \mathbb{R}^N$ is injective.

Proof. The map $\phi^{-1} : V \cap X \rightarrow U$ is a \mathcal{C}^∞ map, so (shrinking V if necessary) we can assume there is a \mathcal{C}^∞ map $\psi : V \rightarrow U$ with $\psi = \phi^{-1}$ on $V \cap X$. Since ϕ maps U onto $V \cap X$, we have $\psi \circ \phi = \phi^{-1} \circ \phi = I =$ the identity map of U onto itself. Thus,

$$I = D(\psi \circ \phi)(0) = (D\psi)(p)(D\phi)(0). \quad (6.22)$$

That is, $D\psi(p)$ is a “left inverse” of $D\phi(0)$. So, $D\phi(0)$ is injective. \square

We can conclude that $\phi : U \rightarrow \mathbb{R}^N$ is an immersion at 0. The canonical immersion theorem tells us that there exists a neighborhood U_0 of 0 in U , a neighborhood V_p of p in V , and a \mathcal{C}^∞ map $g : V_p \rightarrow \mathbb{R}^N$ mapping p onto 0 and mapping V_p diffeomorphically onto a neighborhood \mathcal{O} of 0 in \mathbb{R}^N such that

$$\iota^{-1}(\mathcal{O}) = U_0 \tag{6.23}$$

and

$$g \circ \phi = \iota \tag{6.24}$$

on U_0 . Here, the map ι is the canonical submersion map $\iota : \mathbb{R}^n \rightarrow \mathbb{R}^N$ that maps $(x_1, \dots, x_n) \rightarrow (x_1, \dots, x_n, 0, \dots, 0)$.

By Equation 6.24, the function g maps $\phi(U_0)$ onto $\iota(U_0)$. However, by Equation 6.23, the set $\iota(U_0)$ is the subset of \mathcal{O} defined by the equations

$$x_i = 0, \quad i = n + 1, \dots, N. \tag{6.25}$$

So, if $g = (g_1, \dots, g_N)$, then $\phi(U_0) = X \cap V_p$ is defined by the equations

$$g_i = 0, \quad i = n + 1, \dots, N. \tag{6.26}$$

Moreover, the $N \times N$ matrix

$$\left[\frac{\partial g_i}{\partial x_j}(x) \right] \tag{6.27}$$

is of rank N at every point $x \in V_p$, since $g : V_p \rightarrow \mathcal{O}$ is a diffeomorphism. Hence, the last $N - n$ row vectors of this matrix

$$\left(\frac{\partial g_i}{\partial x_1}, \dots, \frac{\partial g_i}{\partial x_N} \right), \quad i = n + 1, \dots, N, \tag{6.28}$$

are linearly independent at every point $x \in V_p$.

Now let $k = N - n$ and let $f_i = g_{i+n}$, $i = 1, \dots, k$. Then $X \cap V_p$ is defined by the equations

$$f_i(x) = 0, \quad i = 1, \dots, k, \tag{6.29}$$

and the $k \times N$ matrix

$$\left[\frac{\partial f_i}{\partial x_k}(x) \right] \tag{6.30}$$

is of rank k at all points $x \in V_p$. In other words, the system of equations 6.29 is non-degenerate. \square

Lecture 32

6.4 Tangent Spaces of Manifolds

We generalize our earlier discussion of tangent spaces to tangent spaces of manifolds. First we review our earlier treatment of tangent spaces.

Let $p \in \mathbb{R}^n$. We define

$$T_p\mathbb{R}^n = \{(p, v) : v \in \mathbb{R}^n\}. \quad (6.31)$$

Of course, we associate $T_p\mathbb{R}^n \cong \mathbb{R}^n$ by the map $(p, v) \rightarrow v$.

If U is open in \mathbb{R}^n , V is open in \mathbb{R}^k , and $f : (U, p) \rightarrow (V, q)$ (meaning that f maps $U \rightarrow V$ and $p \rightarrow p_1$) is a \mathcal{C}^∞ map, then we have the map $df_p : T_p\mathbb{R}^n \rightarrow T_q\mathbb{R}^k$. Via the identifications $T_p\mathbb{R}^n \cong \mathbb{R}^n$ and $T_q\mathbb{R}^k \cong \mathbb{R}^k$, the map df_p is just the map $Df(p) : \mathbb{R}^n \rightarrow \mathbb{R}^k$. Because these two maps can be identified, we can use the chain rule for \mathcal{C}^∞ maps. Specifically, if $f : (U, p) \rightarrow (V, q)$ and $g : (V, q) \rightarrow (\mathbb{R}^\ell, w)$, then

$$d(g \circ f)_p = (dg)_q \circ (df)_p, \quad (6.32)$$

because $(Dg)(q)(Df(p)) = (Dg \circ f)(p)$.

You might be wondering: Why did we make everything more complicated by using df instead of Df ? The answer is because we are going to generalize from Euclidean space to manifolds.

Remember, a set $X \subseteq \mathbb{R}^N$ is an n -dimensional manifold if for every $p \in X$, there exists a neighborhood V of p in \mathbb{R}^N , an open set U in \mathbb{R}^n , and a diffeomorphism $\phi : U \rightarrow V \cap X$. The map $\phi : U \rightarrow V \cap X$ is called a parameterization of X at p .

Let us think of ϕ as a map $\phi : U \rightarrow \mathbb{R}^N$ with $\text{Im } \phi \subseteq X$.

Claim. *Let $\phi^{-1}(p) = q$. Then the map $(d\phi)_q : T_q\mathbb{R}^n \rightarrow T_p\mathbb{R}^N$ is one-to-one.*

Reminder of proof: The map $\phi^{-1} : V \cap X \rightarrow U$ is a \mathcal{C}^∞ map. So, shrinking V if necessary, we can assume that this map extends to a map $\psi : V \rightarrow U$ such that $\psi = \phi^{-1}$ on $X \cap V$. Then note that for any $u \in U$, we have $\psi(\phi(u)) = \phi^{-1}(\phi(u)) = u$. So, $\psi \circ \phi = \text{id}_U =$ the identity on U .

Using the chain rule, and letting $\phi(q) = p$, we get

$$\begin{aligned} d(\psi \circ \phi)_q &= (d\psi)_o \circ (d\phi)_q \\ &= (d(\text{id}_U))_q. \end{aligned} \quad (6.33)$$

So, $(d\phi)_q$ is injective. □

Today we define for any $p \in X$ the tangent space T_pX , which will be a vector subspace $T_pX \subseteq T_p\mathbb{R}^N$. The tangent space will be like in elementary calculus, that is, a space tangent to some surface.

Let $\phi : U \rightarrow V \cap X$ be a parameterization of X , and let $\phi(q) = p$. The above claim tells us that $(d\phi)_q : T_q\mathbb{R}^n \rightarrow T_p\mathbb{R}^N$ is injective.

Definition 6.10. We define the *tangent space* of a manifold X to be

$$T_pX = \text{Im}(d\phi)_q. \quad (6.34)$$

Because $(d\phi)_q$ is injective, the space T_pX is n -dimensional.

We would like to show that the space T_pX does not depend on the choice of parameterization ϕ . To do so, we will make use of an equivalent definition for the tangent space T_pX .

Last time we showed that given $p \in X \subseteq \mathbb{R}^N$, and $k = N - n$, there exists a neighborhood V of p in \mathbb{R}^N and a \mathcal{C}^∞ map $f : V \rightarrow \mathbb{R}^k$ mapping $f(p) = 0$ such that $X \cap V = f^{-1}(0)$. Note that $f^{-1}(0) \cap C_f = \emptyset$ (where here \emptyset is the empty set).

We motivate the second definition of the tangent space. Since $p \in f^{-1}(0)$, the point $p \notin C_f$. So, the map $df_p : T_p\mathbb{R}^N \rightarrow T_0\mathbb{R}^k$ is surjective. So, the kernel of df_p in $T_p\mathbb{R}^N$ is of dimension $N - k = n$.

Definition 6.11. An alternate definition for the *tangent space* of a manifold is

$$T_pX = \ker df_p. \quad (6.35)$$

Claim. *These two definitions for the tangent space T_pX are equivalent.*

Proof. Let $\phi : U \rightarrow V \cap X$ be a parameterization of X at p with $\phi(p) = q$. The function $f : V \rightarrow \mathbb{R}^k$ has the property that $f^{-1}(0) = X \cap V$. So, $f \circ \phi \equiv 0$. Applying the chain rule,

$$(df_p) \circ (d\phi)_q = d(0) = 0. \quad (6.36)$$

So, $\text{Im } d\phi_q = \ker df_p$. □

We can now explain why the tangent space T_pX is independent of the chosen parameterization. We have two definitions for the tangent space. The first does not depend on the choice of ϕ , and the second does not depend on choice of f . Therefore, the tangent space depends on neither.

Lemma 6.12. *Let W be an open subset of \mathbb{R}^ℓ , and let $g : W \rightarrow \mathbb{R}^n$ be a \mathcal{C}^∞ map. Suppose that $g(W) \subseteq X$ and that $g(w) = p$, where $w \in W$. Then $(dg)_w \subseteq T_pX$.*

Proof Hint: We leave the proof as an exercise. As above, we have a map $f : V \rightarrow \mathbb{R}^k$ such that $X \cap V = f^{-1}(0)$ and $T_pX = \ker df_p$. Let $W_1 = g^{-1}(V)$, and consider the map $f \circ g : W_1 \rightarrow \mathbb{R}^k$. As before, $f \circ g = 0$, so $df_p \circ dg_w = 0$. □

Suppose that $X \subseteq \mathbb{R}^N$ is an n -dimensional manifold and $Y \subseteq \mathbb{R}^\ell$ is an m -dimensional manifold. Let $f : X \rightarrow Y$ be a \mathcal{C}^∞ map, and let $f(p) = q$. We want to define a linear map

$$df_p : T_pX \rightarrow T_qY. \quad (6.37)$$

Let v be a neighborhood of p in \mathbb{R}^N , and let $g : V \rightarrow \mathbb{R}^\ell$ be a map such that $g = f$ on $V \cap X$. By definition $T_pX \subseteq T_p\mathbb{R}^N$, so we have

$$dg_p : T_p\mathbb{R}^N \rightarrow T_q\mathbb{R}^\ell. \quad (6.38)$$

We define the map df_p to be the restriction of dg_p to the tangent space T_pX .

Definition 6.13.

$$df_p = dg_p|_{T_p X}. \quad (6.39)$$

There are two questions about this definition that should have us worried:

1. Is $\text{Im } dg_p(T_p X)$ a subset of $T_q Y$?
2. Does this definition depend on the choice of g ?

We address these two questions here:

1. Is $\text{Im } dg_p(T_p X)$ a subset of $T_q Y$?

Let U be an open subset of \mathbb{R}^N , let $q = f(p)$, and let $\phi : U \rightarrow X \cap V$ be a parameterization of X at p . As before, let us think of ϕ as a map $\phi : U \rightarrow \mathbb{R}^N$ with $\phi(U) \subseteq X$.

By definition, $T_p X = \text{Im } (d\phi)_r$, where $\phi(r) = p$. So, given $v \in T_p X$, one can always find $w \in T_r \mathbb{R}^n$ with $v = (d\phi)_r w$.

Now, is it true that $(dg)_p(v) \in T_q Y$? We have

$$\begin{aligned} (dg)_p v &= (dg)_p (d\phi)_r(w) \\ &= d(g \circ \phi)_r(w), \end{aligned} \quad (6.40)$$

and the map $(g \circ \phi)$ is of the form $g \circ \phi : U \rightarrow Y$, so

$$d(g \circ \phi)_r(w) \in T_q Y. \quad (6.41)$$

2. Does the definition depend on the choice of g ?

Consider two such maps $g_1, g_2 : V \rightarrow \mathbb{R}^\ell$. They satisfy $g_1 = g_2 = f$ on $X \cap V$. Then, with v, w as above,

$$(dg_1)_p(v) = d(g_1 \circ \phi)_r(w) \quad (6.42)$$

$$(dg_2)_p(v) = d(g_2 \circ \phi)_r(w). \quad (6.43)$$

Since $g_1 = g_2$ on $X \cap V$, we have

$$g_1 \circ \phi = g_2 \circ \phi = f \circ \phi. \quad (6.44)$$

Hence,

$$d(g_1 \circ \phi)_r(w) = d(g_2 \circ \phi)_r(w). \quad (6.45)$$

As an exercise, show that the chain rule also generalizes to manifolds as follows: Suppose that X_1, X_2, X_3 are manifolds with $X_i \subseteq \mathbb{R}^{N_i}$, and let $f : X_1 \rightarrow X_2$ and $g : X_2 \rightarrow X_3$ be \mathcal{C}^∞ maps. Let $f(p) = q$ and $g(q) = r$.

Show the following claim.

Claim.

$$d(g \circ f)_p = (dg)_q \circ (df)_p. \quad (6.46)$$

Proof Hint: Let V_1 be a neighborhood of p in \mathbb{R}^{N_1} , and let V_2 be a neighborhood of q in \mathbb{R}^{N_2} . Let $\tilde{f} : V_1 \rightarrow V_2$ be an extension of f to V_1 , and let $\tilde{g} : V_2 \rightarrow \mathbb{R}^{N_3}$ be an extension of g to V_2 .

The chain rule for f, g follows from the chain rule for \tilde{f}, \tilde{g} .

□

Lecture 33

6.5 Differential Forms on Manifolds

Let $U \subseteq \mathbb{R}^n$ be open. By definition, a k -form ω on U is a function which assigns to each point $p \in U$ an element $\omega_p \in \Lambda^k(T_p^*\mathbb{R}^n)$.

We now define the notion of a k -form on a manifold. Let $X \subseteq \mathbb{R}^N$ be an n -dimensional manifold. Then, for $p \in X$, the tangent space $T_pX \subseteq T_p\mathbb{R}^N$.

Definition 6.14. A k -form ω on X is a function on X which assigns to each point $p \in X$ an element $\omega_p \in \Lambda^k((T_pX)^*)$.

Suppose that $f : X \rightarrow \mathbb{R}$ is a C^∞ map, and let $f(p) = a$. Then df_p is of the form

$$df_p : T_pX \rightarrow T_a\mathbb{R} \cong \mathbb{R}. \quad (6.47)$$

We can think of $df_p \in (T_pX)^* = \Lambda^1((T_pX)^*)$. So, we get a one-form df on X which maps each $p \in X$ to df_p .

Now, suppose

$$\mu \text{ is a } k\text{-form on } X, \text{ and} \quad (6.48)$$

$$\nu \text{ is an } \ell\text{-form on } X. \quad (6.49)$$

For $p \in X$, we have

$$\mu_p \in \Lambda^k(T_p^*X) \text{ and} \quad (6.50)$$

$$\nu_p \in \Lambda^\ell(T_p^*X). \quad (6.51)$$

Taking the wedge product,

$$\mu_p \wedge \nu_p \in \Lambda^{k+\ell}(T_p^*X). \quad (6.52)$$

The wedge product $\mu \wedge \nu$ is the $(k + \ell)$ -form mapping $p \in X$ to $\mu_p \wedge \nu_p$.

Now we consider the pullback operation. Let $X \subseteq \mathbb{R}^N$ and $Y \subseteq \mathbb{R}^\ell$ be manifolds, and let $f : X \rightarrow Y$ be a C^∞ map. Let $p \in X$ and $a = f(p)$. We have the map

$$df_p : T_pX \rightarrow T_aY. \quad (6.53)$$

From this we get the pullback

$$(df_p)^* : \Lambda^k(T_a^*Y) \rightarrow \Lambda^k(T_p^*X). \quad (6.54)$$

Let ω be a k -form on Y . Then $f^*\omega$ is defined by

$$(f^*\omega)_p = (df_p)^*\omega_q. \quad (6.55)$$

Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be \mathcal{C}^∞ maps on manifolds X, Y, Z . Let ω be a k -form. Then

$$(g \circ f)^*\omega = f^*(g^*\omega), \quad (6.56)$$

where $g \circ f : X \rightarrow Z$.

So far, the treatment of k -forms for manifolds has been basically the same as our earlier treatment of k -forms. However, the treatment for manifolds becomes more complicated when we study \mathcal{C}^∞ forms.

Let U be an open subset of \mathbb{R}^n , and let ω be a k -form on U . We can write

$$\omega = \sum a_I(x) dx_{i_1} \wedge \cdots \wedge dx_{i_k}, \quad I = (i_1, \dots, i_k). \quad (6.57)$$

By definition, we say that $\omega \in \Omega^k(U)$ if each $A_I \in \mathcal{C}^\infty(U)$.

Let V be an open subset of \mathbb{R}^k , and let $f : U \rightarrow V$ be a \mathcal{C}^∞ map. Let $\omega \in \Omega^k(V)$. Then $f^*\omega \in \Omega^k(U)$. Now, we want to define what we mean by a \mathcal{C}^∞ form on a manifold.

Let $X \subseteq \mathbb{R}^n$ be an n -dimensional manifold, and let $p \in X$. There exists an open set U in \mathbb{R}^n , a neighborhood V of p in \mathbb{R}^n , and a diffeomorphism $\phi : U \rightarrow V \cap X$. The diffeomorphism ϕ is a parameterization of X at p .

We can think of ϕ in the following two ways:

1. as a map of U onto $V \cap X$, or
2. as a map of U onto V , whose image is contained in X .

The second way of thinking about ϕ is actually the map $\iota_X \circ \phi$, where $\iota_X : X \rightarrow \mathbb{R}^n$ is the inclusion map. Note that $\iota_X : X \rightarrow \mathbb{R}^n$ is \mathcal{C}^∞ , because it extends to the identity map $I : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

We give two equivalent definitions for \mathcal{C}^∞ k -forms. Let ω be a k -form on X .

Definition 6.15. The k -form ω is \mathcal{C}^∞ at p if there exists a k -form $\tilde{\omega} \in \Omega^k(V)$ such that $\iota_X^*\tilde{\omega} = \omega$.

Definition 6.16. The k -form ω is \mathcal{C}^∞ at p if there exists a diffeomorphism $\phi : U \rightarrow V \cap X$ such that $\phi^*\omega \in \Omega^k(U)$.

The first definition depends only on the choice of $\tilde{\omega}$, and the second definition depends only on the choice of ϕ . So, if the definitions are equivalent, then neither definition depends on the choice of $\tilde{\omega}$ or the choice of ϕ .

We show that these two definitions are indeed equivalent.

Claim. *The above two definitions are equivalent.*

Proof. First, we show that (def 6.15) \implies (def 6.16). Let $\omega = \iota_X^*\tilde{\omega}$. Then $\phi^*\omega = (\iota_X \circ \phi)^*\tilde{\omega}$. The map $\iota_X \circ \phi : U \rightarrow V$ is \mathcal{C}^∞ , and $\tilde{\omega} \in \Omega^k(V)$, so $\phi^*\omega = (\iota_X \circ \phi)^*\tilde{\omega} \in \Omega^k(U)$.

Second, we show that (def 6.16) \implies (def 6.15). Let $\phi : U \rightarrow V \cap U$ be a diffeomorphism. Then $\phi^{-1} : V \cap X \rightarrow U$ can be extended to $\psi : V \rightarrow U$, where ψ is \mathcal{C}^∞ . On $V \cap X$, the map $\phi = \iota_X^* \tilde{\omega}$, where $\tilde{\omega} = \psi^*(\phi^*\omega)$. It is easy to show that $\tilde{\omega}$ is \mathcal{C}^∞ . \square

Definition 6.17. The k -form ω is \mathcal{C}^∞ if ω is \mathcal{C}^∞ at p for every point $p \in X$.

Notation. If ω is \mathcal{C}^∞ , then $\omega \in \Omega^k(X)$.

Theorem 6.18. If $\omega \in \Omega^k(X)$, then there exists a neighborhood W of X in \mathbb{R}^N and a k -form $\tilde{\omega} \in \Omega^k(W)$ such that $\iota_X^* \tilde{\omega} = \omega$.

Proof. Let $p \in X$. There exists a neighborhood V_p of p in \mathbb{R}^N and a k -form $\omega^p \in \Omega^k(V_p)$ such that $\iota_X^* \omega^p = \omega$ on $V_p \cap X$.

Let

$$W \subseteq \bigcup_{p \in X} V_p. \quad (6.58)$$

The collection of sets $\{V_p : p \in X\}$ is an open cover of W . Let ρ_i , $i = 1, 2, 3, \dots$, be a partition of unity subordinate to this cover. So, $\rho_i \in \mathcal{C}_0^\infty(W)$ and $\text{supp } \rho_i \subset V_p$ for some p . Let

$$\tilde{\omega}_i = \begin{cases} \rho_i \omega^p & \text{on } V_p, \\ 0 & \text{elsewhere.} \end{cases} \quad (6.59)$$

Notice that

$$\begin{aligned} \iota_X^* \tilde{\omega}_i &= \iota_X^* \rho_i \iota_X^* \omega^p \\ &= (\iota_X^* \rho_i) \omega. \end{aligned} \quad (6.60)$$

Take

$$\tilde{\omega} = \sum_{i=1}^{\infty} \tilde{\omega}_i. \quad (6.61)$$

This sum makes sense since we used a partition of unity. From the sum, we can see that $\tilde{\omega} \in \Omega^k(W)$. Finally,

$$\begin{aligned} \iota_X^* \tilde{\omega} &= (\iota_X^* \sum \rho_i) \omega \\ &= \omega. \end{aligned} \quad (6.62)$$

\square

Theorem 6.19. Let $X \subseteq \mathbb{R}^N$ and $Y \subseteq \mathbb{R}^\ell$ be manifolds, and let $f : X \rightarrow Y$ be a \mathcal{C}^∞ map. If $\omega \in \Omega^k(X)$, then $f^*\omega \in \Omega^k(Y)$.

Proof. Take an open set W in \mathbb{R}^ℓ such that $W \supset Y$, and take $\tilde{\omega} \in \Omega^k(W)$ such that $\iota_X^* \tilde{\omega} = \omega$. Take any $p \in X$ and $\phi : U \rightarrow V$ a parameterization of X at p .

We show that the pullback $\phi^*(f^*\omega)$ is in $\Omega^k(U)$. We can write

$$\begin{aligned}\phi^*(f^*\omega) &= \phi^* f^*(\iota_X^* \tilde{\omega}) \\ &= (\iota \circ f \circ \phi)^* \tilde{\omega},\end{aligned}\tag{6.63}$$

where in the last step we used the chain rule.

The form $\tilde{\omega} \in \Omega^k(W)$, where W is open in \mathbb{R}^ℓ , so $\iota \circ f \circ \phi : U \rightarrow W$. The theorem that we proved on Euclidean spaces shows that the r.h.s of Equation 6.63 is in $\Omega^k(U)$. \square

The student should check the following claim:

Claim. If $\mu, \nu \in \Omega^k(Y)$, then

$$f^*(\mu \wedge \nu) = f^*\mu \wedge f^*\nu.\tag{6.64}$$

The differential operation d is an important operator on k -forms on manifolds.

$$d : \Omega^k(X) \rightarrow \Omega^{k+1}(X).\tag{6.65}$$

Let $X \subseteq \mathbb{R}^N$ be a manifold, and let $\omega \in \Omega^k(X)$. There exists an open neighborhood W of X in \mathbb{R}^N and a k -form $\tilde{\omega} \in \Omega^k(W)$ such that $\iota_X^* \tilde{\omega} = \omega$.

Definition 6.20. $d\omega = \iota_X^* d\tilde{\omega}$.

Why is this definition well-defined? It seems to depend on the choice of $\tilde{\omega}$.

Take a parameterization $\phi : U \rightarrow V \cap X$ of X at p . Then

$$\begin{aligned}\phi^* \iota_X^* d\tilde{\omega} &= (\iota_X \circ \phi)^* d\tilde{\omega} \\ &= d(\iota_X \circ \phi)^* \tilde{\omega} \\ &= d\phi^*(\iota_X^* \tilde{\omega}) \\ &= d\phi^*\omega.\end{aligned}\tag{6.66}$$

So,

$$\phi^* \iota_X^* d\tilde{\omega} = d\phi^*\omega.\tag{6.67}$$

Take the inverse mapping $\phi^{-1} : V \cap X \rightarrow U$ and take the pullback $(\phi^{-1})^*$ of each side of Equation 6.67, to obtain

$$\iota_X^* d\tilde{\omega} = (\phi^{-1})^* d\phi^*\omega.\tag{6.68}$$

The r.h.s does not depend on $\tilde{\omega}$, so neither does the l.h.s.

To summarize this lecture, everything we did with k -forms on Euclidean space applies to k -forms on manifolds.

Lecture 34

6.6 Orientation of Manifolds

Let X be an n -dimensional manifold in \mathbb{R}^N . Assume that X is a closed subset of \mathbb{R}^N . Let $f : X \rightarrow \mathbb{R}$ be a C^∞ map.

Definition 6.21. We remind you that the support of f is defined to be

$$\text{supp } f = \overline{\{x \in X : f(x) \neq 0\}}. \quad (6.69)$$

Since X is closed, we don't have to worry about whether we are taking the closure in X or in \mathbb{R}^n .

Note that

$$f \in C_0^\infty(X) \iff \text{supp } f \text{ is compact.} \quad (6.70)$$

Let $\omega \in \Omega^k(X)$. Then

$$\text{supp } \omega = \overline{\{p \in X : \omega_p \neq 0\}}. \quad (6.71)$$

We use the notation

$$\omega \in \Omega_c^k(X) \iff \text{supp } \omega \text{ is compact.} \quad (6.72)$$

We will be using partitions of unity, so we remind you of the definition:

Definition 6.22. A collection of functions $\{\rho_i \in C_0^\infty(X) : i = 1, 2, 3, \dots\}$ is a *partition of unity* if

1. $0 \leq \rho_i$,
2. For every compact set $A \subseteq X$, there exists $N > 0$ such that $\text{supp } \rho_i \cap A = \emptyset$ for all $i > N$,
3. $\sum \rho_i = 1$.

Suppose the collection of sets $\mathcal{U} = \{U_\alpha : \alpha \in I\}$ is a covering of X by open subsets U_α of X .

Definition 6.23. The partition of unity ρ_i , $i = 1, 2, 3, \dots$, is *subordinate to* \mathcal{U} if for every i , there exists $\alpha \in I$ such that $\text{supp } \rho_i \subseteq U_\alpha$.

Claim. Given a collection of sets $\mathcal{U} = \{U_\alpha : \alpha \in I\}$, there exists a partition of unity subordinate to \mathcal{U} .

Proof. For each $\alpha \in I$, let \tilde{U}_α be an open set in \mathbb{R}^N such that $U_\alpha = \tilde{U}_\alpha \cap X$. We define the collection of sets $\tilde{\mathcal{U}} = \{\tilde{U}_\alpha : \alpha \in I\}$. Let

$$\tilde{U} = \bigcup \tilde{U}_\alpha. \quad (6.73)$$

From our study of Euclidean space, we know that there exists a partition of unity $\tilde{\rho}_i \in \mathcal{C}_0^\infty(\tilde{U})$, $i = 1, 2, 3, \dots$, subordinate to $\tilde{\mathcal{U}}$. Let $\iota_X : X \rightarrow \tilde{\mathcal{U}}$ be the inclusion map. Then

$$\rho_i = \tilde{\rho}_i \circ \iota_X = \iota_X^* \tilde{\rho}_i, \quad (6.74)$$

which you should check. □

We review orientations in Euclidean space before generalizing to manifolds. For a more comprehensive review, read section 7 of the Multi-linear Algebra notes.

Suppose \mathbb{L} is a one-dimensional vector space and that $v \in \mathbb{L} - \{0\}$. The set $\mathbb{L} - \{0\}$ has two components:

$$\{\lambda v : \lambda > 0\} \quad \text{and} \quad \{\lambda v : \lambda < 0\}. \quad (6.75)$$

Definition 6.24. An *orientation of \mathbb{L}* is a choice of one of these components.

Notation. We call the preferred component \mathbb{L}_+ (the positive component). We call the other component \mathbb{L}_- (the negative component).

We define a vector v to be *positively oriented* if $v \in \mathbb{L}_+$.

Now, let V be an n -dimensional vector space.

Definition 6.25. 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 $\Lambda^n(V^*)_+$.

Suppose that V_1, V_2 are oriented n -dimensional vector spaces, and let $A : V_1 \rightarrow V_2$ be a bijective linear map.

Definition 6.26. The map A is *orientation preserving* if

$$\omega \in \Lambda^n(V_2)_+ \implies A^*\omega \in \Lambda^n(V_1)_+. \quad (6.76)$$

Suppose that V_3 is also an oriented n -dimensional vector space, and let $B : V_2 \rightarrow V_3$ be a bijective linear map. If A and B are orientation preserving, then BA is also orientation preserving.

Finally, let us generalize the notion of orientation to orientations of manifolds. Let $X \subseteq \mathbb{R}^N$ be an n -dimensional manifold.

Definition 6.27. An *orientation of X* is a function on X which assigns to each point $p \in X$ an orientation of $T_p X$.

We give two examples of orientations of a manifold:

Example 1: Let $\omega \in \Lambda^n(X)$, and suppose that ω is nowhere vanishing. Orient X by assigning to $p \in X$ the orientation of $T_p X$ for which $\omega_p \in \Lambda^n(T_p^* X)_+$.

Example 2: Take $X = U$, an open subset of \mathbb{R}^n , and let

$$\omega = dx_1 \wedge \cdots \wedge dx_n. \quad (6.77)$$

Define an orientation as in the first example. This orientation is called the *standard orientation of U* .

Definition 6.28. An orientation of X is a \mathcal{C}^∞ orientation if for every point $p \in X$, there exists a neighborhood U of p in X and an n -form $\omega \in \Omega^n(U)$ such that for all points $q \in U$, $\omega_q \in \Lambda^n(T_q^* X)_+$.

From now on, we will only consider \mathcal{C}^∞ orientations.

Theorem 6.29. If X is oriented, then there exists $\omega \in \Omega^n(X)$ such that for all $p \in X$, $\omega_p \in \Lambda^n(T_p^* X)_+$.

Proof. For every point $p \in X$, there exists a neighborhood U_p of p and an n -form $\omega^{(p)} \in \Omega^n(U_p)$ such that for all $q \in U_p$, $(\omega^{(p)})_q \in \Lambda^n(T_q^* X)_+$.

Take ρ_i , $i = 1, 2, \dots$, a partition of unity subordinate to $\mathcal{U} = \{U_p : p \in X\}$. For every i , there exists a point p such that $\rho_i \in \mathcal{C}_0^\infty(U_p)$. Let

$$\omega_i = \begin{cases} \rho_i \omega^{(p)} & \text{on } U_p, \\ 0 & \text{on the } X - U_p. \end{cases} \quad (6.78)$$

Since the ρ_i 's are compactly supported, ω_i is a \mathcal{C}^∞ map. Let

$$\omega = \sum \omega_i. \quad (6.79)$$

One can check that ω is positively oriented at every point. □

Definition 6.30. An n -form $\omega \in \Omega^n(X)$ with the property hypothesized in the above theorem is called a *volume form*.

Remark. If ω_1, ω_2 are volume forms, then we can write $\omega_2 = f\omega_1$, for some $f \in \mathcal{C}^\infty(X)$ (where $f \neq 0$ everywhere). In general, $f(p) > 0$ because $(\omega_1)_p, (\omega_2)_p \in \Lambda^n(T_p^* X)_+$. So, if ω_1, ω_2 are volume forms, then $\omega_2 = f\omega_1$, for some $f \in \mathcal{C}^\infty(X)$ such that $f > 0$.

Remark. Problem #6 on the homework asks you to show that if X is orientable and connected, then there are exactly two ways to orient it. This is easily proved using the above Remark.

Suppose that $X \subseteq \mathbb{R}^n$ is a one-dimensional manifold (a “curve”). Then $T_p X$ is one-dimensional. We can find vectors $v, -v \in T_p X$ such that $\|v\| = 1$. An orientation of X is just a choice of v or $-v$.

Now, suppose that X is an $(n - 1)$ -dimensional manifold in \mathbb{R}^n . Define

$$N_p X = \{v \in T_p \mathbb{R}^n : v \perp w \text{ for all } w \in T_p X\}. \quad (6.80)$$

Then $\dim N_p X = 1$, so you can find $v, -v \in N_p X$ such that $\|v\| = 1$. By Exercise #5 in section 4 of the Multi-linear Algebra Notes, an orientation of $T_p X$ is just a choice of v or $-v$.

Suppose X_1, X_2 are oriented n -dimensional manifolds, and let $f : X_1 \rightarrow X_2$ be a diffeomorphism.

Definition 6.31. The map f is *orientation preserving* if for every $p \in X_1$,

$$df_p : T_p X_1 \rightarrow T_q X_2 \quad (6.81)$$

is orientation preserving, where $q = f(p)$.

Remark. Let ω_2 be a volume form on X_2 . Then f is orientation preserving if and only if $f^* \omega_2 = \omega_1$ is a volume form on X_1 .

We look at an example of what it means for a map to be orientation preserving. Let U, V be open sets on \mathbb{R}^n with the standard orientation. Let $f : U \rightarrow V$ be a diffeomorphism. So, by definition, the form

$$dx_1 \wedge \cdots \wedge dx_n \quad (6.82)$$

is a volume form of V . The form

$$f^* dx_1 \wedge \cdots \wedge dx_n = \det \left[\frac{\partial f_i}{\partial x_j} \right] dx_1 \wedge \cdots \wedge dx_n \quad (6.83)$$

is a volume form of U if and only if

$$\det \left[\frac{\partial f_i}{\partial x_j} \right] > 0, \quad (6.84)$$

that is, if and only if f is orientation preserving in our old sense.

Now that we have studied orientations of manifolds, we have all of the ingredients we need to study integration theory for manifolds.

Lecture 35

Before moving on to integration, we make a few more remarks about orientations.

Let X, Y be oriented manifolds. A diffeomorphism $f : X \rightarrow Y$ is orientation preserving if for every $p \in X$, the map

$$df_p : T_p X \rightarrow T_p Y \quad (6.85)$$

is orientation preserving, where $q = f(p)$.

Let V be open in X , let U be open in \mathbb{R}^n , and let $\phi : U \rightarrow V$ be a parameterization.

Definition 6.32. The map ϕ is an *oriented parameterization* if it is orientation preserving.

Suppose ϕ is orientation reversing. Let $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the linear map defined by

$$A(x_1, \dots, x_n) = (-x_1, x_2, \dots, x_n). \quad (6.86)$$

The map A is orientation reversing. Let $U' = A^{-1}(U)$, and define $\phi' = \phi \circ A : U' \rightarrow V$. Both ϕ and A are orientation reversing, so ϕ' is orientation preserving.

Thus, for every point $p \in X$, there exists an oriented parameterization of X at p .

6.7 Integration on Manifolds

Our goal for today is to take any $\omega \in \Omega_c^n(X)$ and define

$$\int_X \omega. \quad (6.87)$$

First, we consider a special case:

Let $\phi : U \rightarrow V$ be an oriented parameterization. Let U be open in \mathbb{R}^n , and let V be open in X . Take any $\omega \in \Omega_c^n(V)$. Then

$$\int_V \omega = \int_U \phi^* \omega, \quad (6.88)$$

where $\phi^* \omega = f(x) dx_1 \wedge \dots \wedge dx_n$, where $f \in C_0^\infty(U)$ and

$$\int_U \phi^* \omega = \int_U f. \quad (6.89)$$

Claim. The above definition for $\int \omega$ does not depend on the choice of oriented parameterization ϕ .

Proof. Let $\phi_i : U_i \rightarrow V$, $i = 1, 2$, be oriented parameterizations. Let $\omega \in \Omega_c^n(V_1 \cap V_2)$. Define

$$U_{1,2} = \phi_1^{-1}(V_1 \cap V_2), \quad (6.90)$$

$$U_{2,1} = \phi_2^{-1}(V_1 \cap V_2), \quad (6.91)$$

which are open sets in \mathbb{R}^n .

Both ϕ_1 and ϕ_2 are diffeomorphisms, and we have the diagram

$$\begin{array}{ccc} V_1 \cap V_2 & \xlongequal{\quad} & V_1 \cap V_2 \\ \phi_1 \uparrow & & \phi_2 \uparrow \\ U_{1,2} & \xrightarrow{f} & U_{2,1}. \end{array} \quad (6.92)$$

Therefore, $f = \phi_2^{-1} \circ \phi_1$ is a diffeomorphism, and $\phi_1 = \phi_2 \circ f$. Integrating,

$$\begin{aligned} \int_{U_1} \phi_1^* \omega &= \int_{U_{1,2}} \phi_1^* \omega \\ &= \int_{U_{1,2}} (\phi_2 \circ f)^* \omega \\ &= \int_{U_{1,2}} f^* (\phi_2^* \omega). \end{aligned} \quad (6.93)$$

Note that f is orientation preserving, because ϕ_1 and ϕ_2 are orientation preserving.

Using the change of variables formula,

$$\begin{aligned} \int_{U_{1,2}} f^* \phi_2^* \omega &= \int_{U_{2,1}} \phi_2^* \omega \\ &= \int_{U_2} \phi_2^* \omega. \end{aligned} \quad (6.94)$$

So, for all $\omega \in \Omega_c^n(V_1 \cap V_2)$,

$$\int_{V_1} \omega = \int_{U_1} \phi_1^* \omega = \int_{U_2} \phi_2^* \omega = \int_{V_2} \omega. \quad (6.95)$$

□

Above, we showed above how to take integrals over open sets, and now we generalize.

To define the integral, we need the following two inputs:

1. a set of oriented parameterizations $\phi_i : U_i \rightarrow V_i$, $i = 1, 2, \dots$, such that $X = \bigcup V_i$,

2. a partition of unity $\rho_i \in \mathcal{C}_0^\infty(V_i)$ subordinate to the cover $\{V_i\}$.

Definition 6.33. Let $\omega \in \Omega_c^n(X)$. We define the integral

$$\int_X \omega = \sum_{i=1}^{\infty} \int_{V_i} \rho_i \omega. \quad (6.96)$$

One can check various standard properties of integrals, such as linearity:

$$\int_X \omega_1 + \omega_2 = \int_X \omega_1 + \int_X \omega_2. \quad (6.97)$$

We now show that this definition is independent of the choice of the two inputs (the parameterizations and the partition of unity).

Consider two different inputs:

1. oriented parameterizations $\phi_j' : U_j' \rightarrow V_j'$, $j = 1, 2, \dots$, such that $X = \bigcup V_j'$,
2. a partition of unity $\rho_i' \in \mathcal{C}_0^\infty(V_j')$ subordinate to the cover $\{V_j'\}$.

Then,

$$\begin{aligned} \int_{V_i} \rho_i \omega &= \int_{V_i} \left(\sum_{j=1}^{\infty} \rho_j' \omega \right) \\ &= \sum_{j=1}^{\infty} \int_{V_i} \rho_i \rho_j' \omega \\ &= \sum_{j=1}^{\infty} \int_{V_i \cap V_j'} \rho_i \rho_j' \omega. \end{aligned} \quad (6.98)$$

Summing over i ,

$$\begin{aligned} \sum_i \int_{V_i} \rho_i \omega &= \sum_{i,j=1}^{\infty} \int_{V_i \cap V_j'} \rho_i \rho_j' \omega \\ &= \sum_j \int_{V_j'} \rho_j' \omega, \end{aligned} \quad (6.99)$$

where the first term equals the last term by symmetry. Therefore, the integral $\int \omega$ is independent of the choices of these two inputs.

Let $X \subseteq \mathbb{R}^N$ be an oriented connected n -dimensional manifold.

Theorem 6.34. For any $\omega \in \Omega_c^n(X)$, the following are equivalent:

1. $\int_X \omega = 0$,

2. $\omega \in d\Omega_c^{n-1}(X)$.

Proof. This will be a five step proof:

Step 1: The following lemma is called the Connectivity Lemma.

Lemma 6.35. *Given $p, q \in X$, there exists open sets W_j , $j = 0, \dots, N+1$, such that each 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 $W_i \cap W_{i+1} \neq \emptyset$.*

Proof Idea: Fix p . The points q for which this is true form an open set. The points q for which this isn't true also form an open set. Since X is connected, only one of these sets is in X . \square

Step 2: Let $\omega_1, \omega_2 \in \Omega_c^n(X)$. We say that $\omega_1 \sim \omega_2$ if

$$\int_X \omega_1 = \int_X \omega_2. \quad (6.100)$$

We can restate the theorem as

$$\omega_1 \sim \omega_2 \iff \omega_1 - \omega_2 \in d\Omega_c^{n-1}(X). \quad (6.101)$$

Step 3: It suffices to prove the statement (6.101) for $\omega_1 \in \Omega_c^n(V)$ and $\omega_2 \in \Omega_c^n(V')$, where V, V' are diffeomorphic to open sets in \mathbb{R}^n .

Step 4: We use a partition of unity

Lemma 6.36. *The theorem is true if $V = V'$.*

Proof. Let $\phi : U \rightarrow V$ be an orientation preserving parameterization. If $\omega_1 \sim \omega_2$, then

$$\int \phi^* \omega_1 = \int \phi^* \omega_2, \quad (6.102)$$

which is the same as saying that

$$\phi^* \omega_1 - \phi^* \omega_2 \in d\Omega_c^{n-1}(U), \quad (6.103)$$

which is the same as saying that

$$\omega_1 - \omega_2 \in d\Omega_c^{n-1}(V). \quad (6.104)$$

\square

Step 5: In general, by the Connectivity Lemma, there exists sets W_i , $i = 0, \dots, N+1$, such that each W_i is diffeomorphic to an open set in \mathbb{R}^n . We can choose $W_0 = V$ and $W_{N+1} = V'$ and $W_i \cap W_{i+1} \neq \emptyset$ (where \emptyset here is the empty set).

We can choose $\mu_i \in \Omega_c^n(W_i \cap W_{i+1})$ such that

$$c = \int_V \omega_1 = \int \mu_0 = \dots = \int \mu_{N+1} = \int_{V'} \omega_2. \quad (6.105)$$

So,

$$\omega_1 \sim \mu_0 \sim \cdots \sim \mu_N \sim \omega_2. \quad (6.106)$$

We know that $\mu_0 - \omega_1 \in d\Omega_c^{n-1}$ and $\omega_2 - \mu_{N+1} \in d\Omega_C^{n-1}$. Also, each difference $\omega_i - \omega_{i+1} \in d\Omega_c^{n-1}$. Therefore, $\omega_1 - \omega_2 \in d\Omega_c^{n-1}$. \square

6.8 Degree on Manifolds

Suppose that X_1, X_2 are oriented n -dimensional manifolds, and let $f : X_1 \rightarrow X_2$ be a proper map (that is, for every compact set $A \subseteq X$, the set pre-image $f^{-1}(A)$ is compact). It follows that if $\omega \in \Omega_c^k(X_2)$, then $f^*\omega \in \Omega_c^k(X_1)$.

Theorem 6.37. *If X_1, X_2 are connected and $f : X_1 \rightarrow X_2$ is a proper C^∞ map, then there exists a topological invariant of f (called the degree of f) written $\deg(f)$ such that for every $\omega \in \Omega_c^n(X_2)$,*

$$\int_{X_1} f^*\omega = \deg(f) \int_{X_2} \omega. \quad (6.107)$$

Proof. The proof is pretty much verbatim of the proof in Euclidean space. \square

Let us look at a special case. Let $\phi_1 : U \rightarrow V$ be an oriented parameterization, and let V_1 be open in X_1 . Let $f : X_1 \rightarrow X_2$ be an oriented diffeomorphism. Define $\phi_2 = f \circ \phi_1$, which is of the form $\phi_2 : U \rightarrow V_2$, where $V_2 = f(V_1)$. Notice that ϕ_2 is an oriented parameterization of V_2 .

Take $\omega \in \Omega_c^n(V_2)$ and compute the integral

$$\begin{aligned} \int_{V_1} f^*\omega &= \int_U \phi_1^* f^*\omega \\ &= \int_U (f \circ \phi_1)^*\omega \\ &= \int_U \phi_2^*\omega. \end{aligned} \quad (6.108)$$

The n -form ω is compactly supported on V_2 , so

$$\begin{aligned} \int_{V_1} f^*\omega &= \int_U \phi_2^*\omega \\ &= \int_{X_2} \omega. \end{aligned} \quad (6.109)$$

On the other hand,

$$\int_{X_1} f^*\omega = \int_{V_1} f^*\omega. \quad (6.110)$$

Combining these results,

$$\int_{X_1} f^*\omega = \int_{X_2} \omega. \quad (6.111)$$

Therefore,

$$\deg(f) = 1. \quad (6.112)$$

So, we have proved the following theorem, which is the Change of Variables theorem for manifolds:

Theorem 6.38. *Let X_1, X_2 be connected oriented n -dimensional manifolds, and let $f : X_1 \rightarrow X_2$ be an orientation preserving diffeomorphism. Then, for all $\omega \in \Omega_c^n(X_2)$,*

$$\int_{X_1} f^*\omega = \int_{X_2} \omega. \quad (6.113)$$

Lecture 36

The first problem on today's homework will be to prove the inverse function theorem for manifolds. Here we state the theorem and provide a sketch of the proof.

Let X, Y be n -dimensional manifolds, and let $f : X \rightarrow Y$ be a C^∞ map with $f(p) = p_1$.

Theorem 6.39. *If $df_p : T_p X \rightarrow T_{p_1} Y$ is bijective, then f maps a neighborhood V of p diffeomorphically onto a neighborhood V_1 of p_1 .*

Sketch of proof: Let $\phi : U \rightarrow V$ be a parameterization of X at p , with $\phi(q) = p$. Similarly, let $\phi_1 : U_1 \rightarrow V_1$ be a parameterization of Y at p_1 , with $\phi_1(q_1) = p_1$.

Show that we can assume that $f : V \rightarrow V_1$ (Hint: if not, replace V by $V \cap f^{-1}(V_1)$).

Show that we have a diagram

$$\begin{array}{ccc} V & \xrightarrow{f} & V_1 \\ \phi \uparrow & & \phi_1 \uparrow \\ U & \xrightarrow{g} & U_1, \end{array} \tag{6.114}$$

which defines g ,

$$g = \phi_1^{-1} \circ f \circ \phi, \tag{6.115}$$

$$g(q) = q_1. \tag{6.116}$$

So,

$$(dg)_q = (d\phi_1)_{q_1}^{-1} \circ df_p \circ (d\phi)_q. \tag{6.117}$$

Note that all three of the linear maps on the r.h.s. are bijective, so $(dg)_q$ is a bijection. Use the Inverse Function Theorem for open sets in \mathbb{R}^n . □

This ends our explanation of the first homework problem.

Last time we showed the following. Let X, Y be n -dimensional manifolds, and let $f : X \rightarrow Y$ be a proper C^∞ map. We can define a topological invariant $\deg(f)$ such that for every $\omega \in \Omega_c^n(Y)$,

$$\int_X f^* \omega = \deg(f) \int_Y \omega. \tag{6.118}$$

There is a recipe for calculating the degree, which we state in the following theorem. We lead into the theorem with the following lemma.

First, remember that we defined the set C_f of critical points of f by

$$p \in C_f \iff df_p : T_p X \rightarrow T_q Y \text{ is not surjective,} \tag{6.119}$$

where $q = f(p)$.

Lemma 6.40. *Suppose that $q \in Y - f(C_f)$. Then $f^{-1}(q)$ is a finite set.*

Proof. Take $p \in f^{-1}(q)$. Since $p \notin C_f$, the map df_p is bijective. The Inverse Function Theorem tells us that f maps a neighborhood U_p of p diffeomorphically onto an open neighborhood of q . So, $U_p \cap f^{-1}(q) = p$.

Next, note that $\{U_p : p \in f^{-1}(q)\}$ is an open covering of $f^{-1}(q)$. Since f is proper, $f^{-1}(q)$ is compact, so there exists a finite subcover U_{p_1}, \dots, U_{p_N} . Therefore, $f^{-1}(q) = \{p_1, \dots, p_N\}$. \square

The following theorem gives a recipe for computing the degree.

Theorem 6.41.

$$\deg(f) = \sum_{i=1}^N \sigma_{p_i}, \quad (6.120)$$

where

$$\sigma_{p_i} = \begin{cases} +1 & \text{if } df_{p_i} : T_{p_i}X \rightarrow T_qY \text{ is orientation preserving,} \\ -1 & \text{if } df_{p_i} : T_{p_i}X \rightarrow T_qY \text{ is orientation reversing,} \end{cases} \quad (6.121)$$

Proof. The proof is basically the same as the proof in Euclidean space. \square

We say that $q \in Y$ is a *regular value* of f if $q \notin f(C_f)$. Do regular values exist? We showed that in the Euclidean case, the set of non-regular values is of measure zero (Sard's Theorem). The following theorem is the analogous theorem for manifolds.

Theorem 6.42. *If $q_0 \in Y$ and W is a neighborhood of q_0 in Y , then $W - f(C_f)$ is non-empty. That is, every neighborhood of q_0 contains a regular value (this is known as the Volume Theorem).*

Proof. We reduce to Sard's Theorem.

The set $f^{-1}(q_0)$ is a compact set, so we can cover $f^{-1}(q_0)$ by open sets $V_i \subset X$, $i = 1, \dots, N$, such that each V_i is diffeomorphic to an open set in \mathbb{R}^n .

Let W be a neighborhood of q_0 in Y . We can assume the following:

1. W is diffeomorphic to an open set in \mathbb{R}^n ,
2. $f^{-1}(W) \subset \bigcup V_i$ (which is Theorem 4.3 in the Supp. Notes),
3. $f(V_i) \subseteq W$ (for, if not, we can replace V_i with $V_i \cap f^{-1}(W)$).

Let U and the sets U_i , $i = 1, \dots, N$, be open sets in \mathbb{R}^n . Let $\phi : U \rightarrow W$ and the maps $\phi_i : U_i \rightarrow V_i$ be diffeomorphisms. We have the following diagram:

$$\begin{array}{ccc} V_i & \xrightarrow{f} & W \\ \phi_i, \cong \uparrow & & \phi, \cong \uparrow \\ U_i & \xrightarrow{g_i} & U, \end{array} \quad (6.122)$$

which define the maps g_i ,

$$g_i = \phi^{-1} \circ f \circ \phi_i. \quad (6.123)$$

By the chain rule, $x \in C_{g_i} \implies \phi_i(x) \in C_f$, so

$$\phi_i(C_{g_i}) = C_f \cap V_i. \quad (6.124)$$

So,

$$\phi(g_i(C_{g_i})) = f(C_f \cap V_i). \quad (6.125)$$

Then,

$$f(C_f) \cap W = \bigcup_i \phi(g_i(C_{g_i})). \quad (6.126)$$

Sard's Theorem tells us that $g_i(C_{g_i})$ is a set of measure zero in U , so

$$U - \bigcup_i g_i(C_{g_i}) \text{ is non-empty, so} \quad (6.127)$$

$$W - f(C_f) \text{ is also non-empty.} \quad (6.128)$$

In fact, this set is not only non-empty, but is a very, very “full” set. \square

Let $f_0, f_1 : X \rightarrow Y$ be proper \mathcal{C}^∞ maps. Suppose there exists a proper \mathcal{C}^∞ map $F : X \times [0, 1] \rightarrow Y$ such that $F(x, 0) = f_0(x)$ and $F(x, 1) = f_1(x)$. Then

$$\deg(f_0) = \deg(f_1). \quad (6.129)$$

In other words, the degree is a homotopy. The proof of this is essentially the same as before.

6.9 Hopf Theorem

The Hopf Theorem is a nice application of the homotopy invariance of the degree.

Define the n -sphere

$$S^n = \{v \in \mathbb{R}^{n+1} : \|v\| = 1\}. \quad (6.130)$$

Hopf Theorem. *Let n be even. Let $f : S^n \rightarrow \mathbb{R}^{n+1}$ be a \mathcal{C}^∞ map. Then, for some $v \in S^n$,*

$$f(v) = \lambda v, \quad (6.131)$$

for some scalar $\lambda \in \mathbb{R}$.

Proof. We prove the contrapositive. Assume that no such v exists, and take $w = f(v)$. Consider $w - \langle v, w \rangle v \equiv w - w_1$. It follows that $w - w_1 \neq 0$.

Define a new map $\tilde{f} : S^n \rightarrow S^n$ by

$$\tilde{f}(v) = \frac{f(v) - \langle v, f(v) \rangle v}{\|f(v) - \langle v, f(v) \rangle v\|} \quad (6.132)$$

Note that $(w - w_1) \perp v$, so $\tilde{f}(v) \perp v$.

Define a family of functions

$$f_t : S^n \rightarrow S^n, \tag{6.133}$$

$$f_t(v) = (\cos t)v + (\sin t)\tilde{w}, \tag{6.134}$$

where $\tilde{w} = \tilde{f}(v)$ has the properties $\|\tilde{w}\| = 1$ and $\tilde{w} \perp v$.

We compute the degree of f_t . When $t = 0$, $f_t = \text{id}$, so

$$\deg(f_t) = \deg(f_0) = 1. \tag{6.135}$$

When $t = \pi$, $f_t(v) = -v$. But, if n is even, a map from $S^n \rightarrow S^n$ mapping $v \rightarrow (-v)$ has degree -1 . We have arrived at a contradiction. \square

Lecture 37

6.10 Integration on Smooth Domains

Let X be an oriented n -dimensional manifold, and let $\omega \in \Omega_c^n(X)$. We defined the integral

$$\int_X \omega, \quad (6.136)$$

but we can generalize the integral

$$\int_D \omega, \quad (6.137)$$

for some subsets $D \subseteq X$. We generalize, but only to very simple subsets called *smooth domains* (essentially manifolds-with-boundary). The prototypical smooth domain is the half plane:

$$\mathbb{H}^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 \leq 0\}. \quad (6.138)$$

Note that the boundary of the half plane is

$$\text{Bd}(\mathbb{H}^n) = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 = 0\}. \quad (6.139)$$

Definition 6.43. A closed subset $D \subseteq X$ is a *smooth domain* if for every point $p \in \text{Bd}(D)$, there exists a parameterization $\phi : U \rightarrow V$ of X at p such that $\phi(U \cap \mathbb{H}^n) = V \cap D$.

Definition 6.44. The map ϕ is a *parameterization of D at p* .

Note that $\phi : U \cap \mathbb{H}^n \rightarrow V \cap D$ is a homeomorphism, so it maps boundary points to boundary points. So, it maps $U^b = U \cap \text{Bd}(\mathbb{H}^n)$ onto $V^b = V \cap \text{Bd}(D)$.

Let $\psi = \phi|_{U^b}$. Then $\psi : U^b \rightarrow V^b$ is a diffeomorphism. The set U^b is an open set in \mathbb{R}^{n-1} , and ψ is a parameterization of the $\text{Bd}(D)$ at p . We conclude that

$$\text{Bd}(D) \text{ is an } (n-1)\text{-dimensional manifold.} \quad (6.140)$$

Here are some examples of how smooth domains appear in nature:

Let $f : X \rightarrow \mathbb{R}$ be a \mathcal{C}^∞ map, and assume that $f^{-1}(0) \cap C_f = \emptyset$ (the empty set). That is, for all $p \in f^{-1}(0)$, $df_p \neq 0$.

Claim. *The set $D = \{x \in X : f(x) \leq 0\}$ is a smooth domain.*

Proof. Take $p \in \text{Bd}(D)$, so $p = f^{-1}(0)$. Let $\phi : U \rightarrow V$ be a parameterization of X at p . Consider the map $g = f \circ \phi : U \rightarrow \mathbb{R}$. Let $q \in U$ and $p = \phi(q)$. Then

$$(dg_q) = df_p \circ (d\phi)_q. \quad (6.141)$$

We conclude that $dg_q \neq 0$.

By the canonical submersion theorem, there exists a diffeomorphism ψ such that $g \circ \psi = \pi$, where π is the canonical submersion mapping $(x, \dots, x_n) \rightarrow x_1$. We can write simply $g \circ \psi = x_1$. Replacing $\phi = \phi_{\text{old}}$ by $\phi = \phi_{\text{new}} = \phi_{\text{old}} \circ \psi$, we get the new map $\phi : U \rightarrow V$ which is a parameterization of X at p with the property that $f \circ \phi(x_1, \dots, x_n) = x_1$. Thus, ϕ maps $\mathbb{H}^n \cap U$ onto $D \cap V$. \square

We give an example of using the above claim to construct a smooth domain. Let $X = \mathbb{R}^n$, and define

$$f(x) = 1 - (x_1^2 + \dots + x_n^2). \quad (6.142)$$

By definition,

$$f(x) \leq 0 \iff x \in B^n, \quad (6.143)$$

where $B^n = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ is the “unit ball.” So, the unit ball B^n is a smooth domain.

We now define orientations of smooth domains. Assume that X is oriented, and let D be a smooth domain. Let $\phi : U \rightarrow V$ be a parameterization of D at p .

Definition 6.45. The map ϕ is an *oriented parameterization of D* if it is an oriented parameterization of X .

Assume that $\dim X = n > 1$. We show that you can always find an oriented parameterization.

Let $\phi : U \rightarrow V$ be a parameterization of D at p . Suppose that ϕ is *not* oriented. That is, as a diffeomorphism ϕ is orientation reversing. Let $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the map

$$A(x_1, \dots, x_n) = (x_1, \dots, x_{n-1}, -x_n). \quad (6.144)$$

Then A maps $\mathbb{H}^n \rightarrow \mathbb{H}^n$, and $\phi \circ A$ is orientation preserving. So, $\phi \circ A$ is an oriented parameterization of D at p .

Now, let $\phi : U \rightarrow V$ be an oriented parameterization of D at p . We define

$$U^b = U \cap \text{Bd}(\mathbb{H}^n), \quad (6.145)$$

$$V^b = V \cap \text{Bd}(D), \quad (6.146)$$

$$\psi = \phi|_{U^b}, \quad (6.147)$$

where ψ is a parameterization of $\text{Bd}(D)$ at p .

We oriented $\text{Bd}(D)$ at p by requiring ψ to be an oriented parameterization. We need to check the following claim.

Claim. *The definition of oriented does not depend on the choice of parameterization.*

Proof. Let $\phi_i : U_i \rightarrow V_i$, $i = 1, 2$, be oriented parameterizations of D at p . Define

$$U_{1,2} = \phi_1^{-1}(V_1 \cap V_2), \quad (6.148)$$

$$U_{2,1} = \phi_2^{-1}(V_1 \cap V_2), \quad (6.149)$$

from which we obtain the following diagram:

$$\begin{array}{ccc}
 V_1 \cap V_2 & \xlongequal{\quad} & V_1 \cap V_2 \\
 \phi_1 \uparrow & & \phi_2 \uparrow \\
 U_{1,2} & \xrightarrow{g} & U_{2,1},
 \end{array} \tag{6.150}$$

which defines a map g . By the properties of the other maps ϕ_1, ϕ_2 , the map g is an orientation preserving diffeomorphism of $U_{1,2}$ onto $U_{2,1}$. Moreover, g maps

$$U_{1,2}^b = \text{Bd}(\mathbb{H}^n) \cap U_{1,2} \tag{6.151}$$

onto

$$U_{2,1}^b = \text{Bd}(\mathbb{H}^n) \cap U_{2,1}. \tag{6.152}$$

Let $h = g|_{U_{1,2}^b}$, so $h : U_{1,2}^b \rightarrow U_{2,1}^b$. We want to show that h is orientation preserving. To show this, we write g and h in terms of coordinates.

$$g = (g_1, \dots, g_n), \quad \text{where } g_i = g_i(x_1, \dots, x_n). \tag{6.153}$$

So,

$$g \text{ maps } \mathbb{H}^n \text{ to } \mathbb{H}^n \iff \begin{cases} g_1(x_1, \dots, x_n) < 0 & \text{if } x_1 < 0, \\ g_1(x_1, \dots, x_n) > 0 & \text{if } x_1 > 0, \\ g_1(0, x_2, \dots, x_n) = 0 \end{cases} \tag{6.154}$$

These conditions imply that

$$\begin{cases} \frac{\partial}{\partial x_1} g_1(0, x_2, \dots, x_n) \geq 0, \\ \frac{\partial}{\partial x_i} g_1(0, x_2, \dots, x_n) = 0, \text{ for } i \neq 1. \end{cases} \tag{6.155}$$

The map h in coordinates is then

$$\begin{aligned}
 h &= h(x_2, \dots, x_n) \\
 &= (g(0, x_2, \dots, x_n), \dots, g_{n-1}(0, x_2, \dots, x_n)),
 \end{aligned} \tag{6.156}$$

which is the statement that $h = g|_{\text{Bd}(\mathbb{H}^n)}$.

At the point $(0, x_2, \dots, x_n) \in U_{1,2}^b$,

$$Dg = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & 0 & \cdots & 0 \\ * & & & \\ \vdots & Dh & & \\ * & & & \end{bmatrix}. \tag{6.157}$$

The matrix Dg is an $n \times n$ block matrix containing the $(n-1) \times (n-1)$ matrix Dh , because

$$\frac{\partial h_i}{\partial x_j} = \frac{\partial g_i}{\partial x_j}(0, x_2, \dots, x_n), \quad i, j > 1. \tag{6.158}$$

Note that

$$\det(Dg) = \frac{\partial g_1}{\partial x_1} \det(Dh). \quad (6.159)$$

We know that the l.h.s > 0 and that $\frac{\partial g_1}{\partial x_1} > 0$, so $\det(Dh) > 0$. Thus, the map $h : U_{1,2}^b \rightarrow U_{2,1}^b$ is orientation preserving.

To repeat, we showed that in the following diagram, the map h is orientation preserving:

$$\begin{array}{ccc} V_1 \cap V_2 \cap \text{Bd}(D) & \xlongequal{\quad} & V_1 \cap V_2 \cap \text{Bd}(D) \\ \psi_1 \uparrow & & \psi_2 \uparrow \\ U_{1,2}^b & \xrightarrow{\quad h \quad} & U_{2,1}^b. \end{array} \quad (6.160)$$

We conclude that ψ_1 is orientation preserving if and only if ψ_2 is orientation preserving. \square

Lecture 38

We begin with a review from last time.

Let X be an oriented manifold, and let $D \subseteq X$ be a smooth domain. Then $\text{Bd}(D) = Y$ is an oriented $(n - 1)$ -dimensional manifold.

We defined integration over D as follows. For $\omega \in \Omega_c^n(X)$ we want to make sense of the integral

$$\int_D \omega. \tag{6.161}$$

We look at some special cases:

Case 1: Let $p \in \text{Int } D$, and let $\phi : U \rightarrow V$ be an oriented parameterization of X at p , where $V \subseteq \text{Int } D$. For $\omega \in \Omega_c^n(X)$, we define

$$\int_D \omega = \int_V \omega = \int_U \phi^* \omega = \int_{\mathbb{R}^n} \phi^* \omega. \tag{6.162}$$

This is just our old definition for

$$\int_V \omega. \tag{6.163}$$

Case 2: Let $p \in \text{Bd}(D)$, and let $\phi : U \rightarrow V$ be an oriented parameterization of D at p . That is, ϕ maps $U \cap \mathbb{H}^n$ onto $V \cap D$. For $\omega \in \Omega_c^n(V)$, we define

$$\int_D \omega = \int_{\mathbb{H}^n} \phi^* \omega. \tag{6.164}$$

We showed last time that this definition does not depend on the choice of parameterization.

General case: For each $p \in \text{Int } D$, let $\phi : U_p \rightarrow V_p$ be an oriented parameterization of X at p with $V_p \subseteq \text{Int } D$. For each $p \in \text{Bd}(D)$, let $\phi : U_p \rightarrow V_p$ be an oriented parameterization of D at p . Let

$$U = \sum_{p \in D} U_p, \tag{6.165}$$

where the set $\mathcal{U} = \{U_p : p \in D\}$ be an an open cover of U . Let ρ_i , $i = 1, 2, \dots$, be a partition of unity subordinate to this cover.

Definition 6.46. For $\omega \in \Omega_c^n(X)$ we define the integral

$$\int_D \omega = \sum_i \int_D \rho_i \omega. \tag{6.166}$$

Claim. *The r.h.s. of this definition is well-defined.*

Proof. Since the ρ_i 's are a partition of unity, there exists an N such that

$$\text{supp } \omega \cap \text{supp } \rho_i = \phi, \quad (6.167)$$

for all $i > N$.

Hence, there are only a finite number of non-zero terms in the summand. Moreover, each summand is an integral of one of the two types above (cases 1 and 2), and is therefore well-defined. \square

Claim. *The l.h.s. of the definition does not depend on the choice of the partition of unity ρ_i .*

Proof. We proved an analogous assertion about the definition of $\int_X \omega$ a few lectures ago, and the proof of the present claim is exactly the same. \square

6.11 Stokes' Theorem

Stokes' Theorem. *For all $\omega \in \Omega_c^{n-1}(X)$,*

$$\int_D d\omega = \int_{\text{Bd}(D)} \omega. \quad (6.168)$$

Proof. Let ρ_i , $i = 1, 2, \dots$, be a partition of unity as defined above. Replacing ω with $\sum \rho_i \omega$, it suffices to prove this for the two special cases below:

Case 1: Let $p \in \text{Int } D$, and let $\phi : U \rightarrow V$ be an oriented parameterization of X at p with $V \subseteq \text{Int } D$. If $\omega \in \Omega_c^{n-1}(V)$, then

$$\int_D d\omega = \int_{\mathbb{R}^n} \phi^* d\omega = \int_{\mathbb{R}^n} d\phi^* \omega = 0. \quad (6.169)$$

Case 2: Let $p \in \text{Bd}(D)$, and let $\phi : U \rightarrow V$ be an oriented parameterization of D at p . Let $U^b = U \cap \text{Bd}(\mathbb{H}^n)$, and let $V^b = V \cap \text{Bd}(D)$. Define $\psi : U^b \rightarrow V^b$, so $\psi : U^b \rightarrow V^b$ is an oriented parameterization of $\text{Bd}(D)$ at p . If $\omega \in \Omega_c^{n-1}(V)$, then

$$\phi^* \omega = \sum f_i(x_1, \dots, x_n) dx_1 \wedge \cdots \wedge \widehat{dx}_i \wedge \cdots \wedge dx_n. \quad (6.170)$$

What is $\psi^* \omega$? Let $\iota : \mathbb{R}^{n-1} \rightarrow \mathbb{R}^n$ be the inclusion map mapping $\text{Bd}(\mathbb{H}^n) \rightarrow \mathbb{R}^n$. The inclusion map ι maps $(x_2, \dots, x_n) \rightarrow (0, x_2, \dots, x_n)$. Then $\phi \circ \iota = \psi$, so

$$\begin{aligned} \psi^* \omega &= \iota^* \phi^* \omega \\ &= \iota^* \left(\sum_{i=1}^n f_i dx_1 \wedge \cdots \wedge \widehat{dx}_i \wedge \cdots \wedge dx_n \right). \end{aligned} \quad (6.171)$$

But,

$$\iota^* dx_1 = d\iota^* x_1 = 0, \quad \text{since } \iota^* x_1 = 0. \quad (6.172)$$

So,

$$\begin{aligned}\psi^*\omega &= \iota^* f_1 dx_2 \wedge \cdots \wedge dx_n \\ &= f_1(0, x_2, \dots, x_n) dx_2 \wedge \cdots \wedge dx_n.\end{aligned}\tag{6.173}$$

Thus,

$$\int_{\text{Bd}(D)} \omega = \int_{\mathbb{R}^{n-1}} \psi^*\omega = \int_{\mathbb{R}^{n-1}} f_1(0, x_2, \dots, x_n) dx_2 \dots dx_n.\tag{6.174}$$

On the other hand,

$$\int_D d\omega = \int_{\mathbb{H}^n} \phi^* d\omega = \int_{\mathbb{H}^n} d\phi^*\omega.\tag{6.175}$$

One should check that

$$\begin{aligned}d\phi^*\omega &= d\left(\sum f_i dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n\right) \\ &= \left(\sum (-1)^{i-1} \frac{\partial f_i}{\partial x_i}\right) dx_1 \wedge \cdots \wedge dx_n.\end{aligned}\tag{6.176}$$

So, each summand

$$\int \frac{\partial f_i}{\partial x_i} dx_1 \dots dx_n\tag{6.177}$$

can be integrated by parts, integrating first w.r.t. the i th variable. For $i > 1$, this is the integral

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{\partial f_i}{\partial x_i} dx_i &= f_i(x_1, \dots, x_n) \Big|_{x_i=-\infty}^{x_i=\infty} \\ &= 0.\end{aligned}\tag{6.178}$$

For $i = 1$, this is the integral

$$\int_{-\infty}^{\infty} \frac{\partial f_1}{\partial x_1}(x_1, \dots, x_n) dx_1 = f_1(0, x_2, \dots, x_n).\tag{6.179}$$

Thus, the total integral of $\phi^* d\omega$ over \mathbb{H}^n is

$$\int f_1(0, x_2, \dots, x_n) dx_2 \dots dx_n.\tag{6.180}$$

We conclude that

$$\int_D d\omega = \int_{\text{Bd}(D)} \omega.\tag{6.181}$$

□

We look at some applications of Stokes' Theorem.

Let D be a smooth domain. Assume that D is compact and oriented, and let $Y = \text{Bd}(D)$. Let Z be an oriented n -manifold, and let $f : Y \rightarrow Z$ be a \mathcal{C}^∞ map.

Theorem 6.47. *If f extends to a \mathcal{C}^∞ map $F : D \rightarrow Z$, then*

$$\deg(f) = 0.\tag{6.182}$$

Corollary 9. *The Brouwer fixed point theorem follows from the above theorem.*