# CALCULUS ON MANIFOLDS: INTEGRATION

JACK CERONI

Contents



*Date*: December 2021.

## 1. Introduction

The goal of these notes is two-fold: to provide an exposition to the general theory of integration as outlined in Spivak and Munkres' books on analysis, and to help myself review for a MAT257 term test at the University of Toronto.

In these notes, I will attempt to re-derive and explain everything covered in the sections in Spivak's text on integration, as well as the proofs outlined in class/problem sets relating to the following list of topics:

- Partitions, upper and lower sums, the "old-technology" definition of the integral, as well as some basic results related to its properties.
- Measure zero, content zero, and volume
- Fubini's theorem
- 

## 2. Basic Theory of Integration

## 3. Necessary and Sufficient Conditions for Integrability

**Lemma 1.**

We now prove a **sexy** theorem.

**Theorem 1.** *A function is integrable if and only if the set of discontinuities of the function is measure-*0*.*

*Proof.* Let us suppose first that $f$ is integrable. Let $B_n$ be the set of points for which the oscillation of $f$ at $x$ is greater than or equal to $1/n$. Let $P$ be a partition of $R$ such that:

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

for some $\epsilon > 0$. Let $P'$ be the collection of all $S \in P$ that intersect $B_n$. Clearly, this is a finite collection of rectangles covering $B_n$. We have:

$$\sum_{S \in P'} \mathrm{vol}(S) \leq \frac{1}{n} \sum_{S \in P'} [M(f,S) - m(f,S)]\mathrm{vol}(S) \leq \sum_{S \in P} [M(f,S) - m(f,S)]\mathrm{vol}(S) < \frac{\epsilon}{n}$$

so it follows that $B_n$ is measure-0. It follows that the union of all $B_n$ is measure-0, and it equal to the set of discontinuities, as this will be precisely the set of all points at which the oscillation of $f$ is non-zero.

Conversely, suppose that $B$ (the set of discontinuities) is measure-0. Choose some $\epsilon > 0$, and recall that $B_\epsilon = \{x \mid f(x) \geq \epsilon\}$ is compact (this was proved above), so it is content-0.

Let $P'$ be a finite collection of rectangles that cover $B_\epsilon$ with volume less than $\epsilon$. Let $M$ be the distance between the maximum and minimum of $f$. If we choose a partition $P$ that refines $P'$, we have:

$$U(f,P) - L(f,P) = \sum_{S \notin P'} [M(f,S) - m(f,S)]\text{vol}(S) + \sum_{S \in P'}$$

□

## 4. Integrating Over Non-Rectangles

Oftentimes, we want to integrate over regions that are not rectangles. In fact, interpreting the integral as the area under some curve, this is how we will determine the "volumes" of regions that are not rectanglular.

To begin, we define the notion of the **indicator function**:

**Definition 1.** *Let $A \subset \mathbb{R}^n$. We define $\chi_A : \mathbb{R}^n \to \mathbb{R}$ as:*

$$\chi_A(x) = \begin{cases} 1 & \textit{if } x \in A \\ 0 & \textit{otherwise} \end{cases}$$

This function tells us whether we are inside some region $A$ or not. Naturally, we define $\text{vol}(A) = \int \chi_A$. However, we still need to:

(1) Decide which region over which we are integrating.
(2) Decide whether we can even evaluate the above integral (in other words, if $\chi_A$ is integrable).

The first of these questions is easy to answer: the integrals of $\chi_A$ over any rectangle containing $A$ are equal (we will check this later). However, before we discuss this, we need to determine when $\chi_A$ is integrable. We have the following theorem:

**Theorem 2.** *Let $R$ be a rectangle containing $A$. Then $\chi_A$ is integrable over $R$ if and only if $A$ is Jordan-measureable. That is, $A$ is bounded and $\partial A$ is measure-0.*

*Proof.* We prove that the set of discontinuities of $\chi_A$, which we call $C$, is equal to $\partial A$. Given $x \in \partial A$, note that for any open rectangle $U$ containing $x$, $U$ will intersect both $A$ and $A^C$, so $\chi_A$ will take on values 1 and 0 on the rectangle. Thus, $\chi_A$ can't be continuous at $x$, by definition.

Conversely, suppose $x$ is a point at which $\chi_A$ is continuous. If $x \in \text{Int}(A)$ or $x \in \text{Ext}(A)$, there will exist some neighbourhood of $x$ ion either of these sets, and clearly, on this set, $\chi_A$ is constant, and $\chi_A$ is continuous at $x$. Thus, $x \in \partial A$.

It follows that $C = \partial A$.

Now, suppose $\chi_A$ is integrable. It follows that the set of discontinuities of $\chi_A$ is measure-0, so $\partial A$ is measure-0. Suppose $\partial A$ is measure-0. Then the set of discontinuities of $\chi_A$ is measure-0, so $\chi_A$ is integrable. □

## 5. Fubini's Theorem

## 6. Partitions of Unity

---

Now, we come to the topic of partitions of unity, which will help us to define a more generalized notion of integration.

We define a partition of unity for some subset $A \subset \mathbb{R}^n$, suboordinate to an open cover $\mathcal{O}$ of $A$ as some collection of $C^\infty$ functions defined on some open set $U$ containing $A$, such that:

(1) For each $x \in A$, we have $0 \leq \phi(x) \leq 1$, for all $\phi$ in the collection.
(2) For each $x \in A$, there is an open $V$ containing $x$ such that all but finitely many $\phi$ are 0 on $V$.
(3) For each $x \in A$, we have $\sum \phi(x) = 1$.
(4) For each $\phi$, there is an open set $U$ in $\mathcal{O}$ such that $\phi = 0$ outside some **closed** set $C$ in $U$. Actually, we can strengthen this by saying that $C$ is **compact**, not just closed.

Before we jump into the proof of existence for these PO1s, we will motivate them, and do some background work which will be useful to us going forward.

---

Partitions of unity essentially allow us to break a function down into a bunch of local components. We know how to integrate (using our old-technology integration) over "local functions", so we can then just use that theory here. Our hope is that if our non-local function behaves nicely (goes to 0 fast enough as it goes outward to infinity), then taking the infinite sum of all these local contributions will be well-defined, and also behave nicely. We take this to be our **new-technology integration**.

---

In order to construct PO1s, we need to decide first how to break functions down. Given a function $f$, and a function $\phi$ that is only non-zero in some small region $R$, then the function $f \cdot \phi$ will also only be non-zero in a small region, so this is the approach we take with breaking down $f$ (as was explained above).

To prove the PO1 theorem, we first want to construct "flat-topped mountain functions", which effectively look like smooth bumps, where the portion where the bump is equal to 1 lies inside a compact region, and the bump is supported in an open set $U$ containing that compact region.

(1) First, we define a useful function:

$$\sigma(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \exp(-1/x) & \text{for } x > 0 \end{cases}$$

One can check that this function is smooth (although at takes a little bit of work).

(2) We claim that there exist smooth 1D bumps. In other words, we want to define a function $\beta_\epsilon \in C^\infty$ such that $\beta_\epsilon \geq 0$, $\beta_\epsilon(0) > 0$, and outside of $(-\epsilon, \epsilon)$, $\beta_\epsilon$ is 0. This is easy:

$$\beta_\epsilon(x) = \sigma(x + \epsilon)\sigma(\epsilon - x)$$

(3) Now, we generalize this to $n$-dimensions, for a bump centred at $a$. This is also easy: we simply define

$$\boldsymbol{\beta}_{\epsilon,\boldsymbol{a}}(\boldsymbol{x}) = \beta_{\epsilon^2}(|\boldsymbol{x} - \boldsymbol{a}|^2)$$

This is a composition of smooth functions, so it is smooth.

(4) Now, we can move onto the final part of the proof. We claim that there exist "smooth step functions" $\theta$, of the form:

$$\theta(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x \geq 1 \end{cases}$$

and in the interval $(0,1)$, $\theta$ can be anything smooth. Indeed, we simply integrate the bump functions that we found previously and normalize:

$$\theta_0(x) = \int_0^x dt\ \beta_{1/2,1/2}(t) \Rightarrow \theta(x) = \frac{\theta_0(x)}{\theta_0(1)}$$

Note that these are the 1D versions of the bump functions.

(5) Now, we actually arrive at the final part of the proof. We claim that the flat-topped mountain that we are looking for is constructed by summing together a bunch of the $\boldsymbol{\beta}$, and shaving their tops to equal 1 using the $\theta$ constructed previously.

Much like in a neural network, $\theta$ acts like an activation function, which sends a sum of $\boldsymbol{\beta}$ to 1 if its value is greater than some threshold. In this case, we can find this lower threshold on compact $C$, as the sum of $\boldsymbol{\beta}$ that we construct is a continuous function on a compact set, so extreme value theorem applies.

For each $x \in C$, find some $\epsilon_x$ such that $\overline{B_{\epsilon_x}} \subset U$. Clearly, the set of $B_{\epsilon_x}$ is an open cover for $C$, so we can find a finite subcover $B_{\epsilon_1},\ ...,\ B_{\epsilon_n}$. Define $f_0$ as:

$$f_0(x) = \sum_{j=1}^{n} \boldsymbol{\beta}_{\epsilon_j, x_j}(x)$$

Clearly, by definition, $f_0$ will be positive on $C$. Thus, it is bounded below by positive $b$, so we let $f(x) = \theta\left(\frac{f_0(x)}{b}\right)$. Clearly, $f$ is smooth, it is 1 on $C$, and it is 0 outside of $U$.

Thus, we have our desired flat-topped mountains.

---

Before we actually jump into the proof of the PO1 theorem, we need to prove one more preliminary. Basically, we want to show that given open $U$ and compact $C$ in $U$, we can find compact $D$ such that $C$ is in $\mathrm{Int}(D)$, and $D$ is in $U$.

This is pretty easy. For each $x \in C$, pick some $\epsilon_x$ such that $\overline{B_{\epsilon_x}} \subset U$. Take a finite subcover $B_{\epsilon_1},\ ...,\ B_{\epsilon_n}$ of $C$, and let $D$ be the union of the closures of these open sets. Clearly, it will be closed and bounded, so it is compact. Its interior contains $C$, and it is in $U$.

---

First, we prove the PO1 theorem in the case that $A$ is compact. The idea is to take an open cover, and then shrink the open cover down to a collection of compact sets which also cover $A$. For each of these compact sets, we use the above constructions to define a bump function $\boldsymbol{\beta}$ on each, and then appropriately re-normalize.

Let's jump into it.

Suppose $A$ is compact, and $\mathcal{O}$ is some open cover. It follows that there exists some finite subcover of $U_i \in \mathcal{O}$. Thus, all we have to do is construct a PO1 that is suboordinate to the cover $\{U_1,\ ...,\ U_n\}$, and we will have an open cover suboordinate to $\mathcal{O}$.

We construct the compact supports in each $U_i$ in an iterative manner. Suppose we have a collection of compact sets $D_1,\ ...,\ D_{k-1}$ such that $D_j \subset U_j$, and $\mathrm{Int}(D_1),\ ...,\ \mathrm{Int}(D_{k-1}), U_k,\ ...,\ U_n$ covers $A$. Then we define $D_k$ by noting that $A$ minus all the sets in the above collection except for $U_k$. This set $S$ will be closed and bounded, so it is compact.

We then use the above preliminary to choose compact $D_k$ such that $S \subset \text{Int}(D_k)$ and $D_k \subset U_k$, as $S$ is clearly a subset of $U_k$. Continuing along with this iterative process gives us our set of desired supports.

---

With this fact, we use the first preliminary to choose smooth $\psi_j$ such that $\psi_j$ is 1 on $D_j$, and 0 outside a closed set in $U_j$. We define functions on the open union of all $U_j$ as:

$$\phi_j^0(x) = \frac{\psi_j(x)}{\psi_1(x) + \cdots + \psi_n(x)}$$

Clearly, this set of functions will satisfy the following criteria:

(1) Each function is smooth.
(2) Each function has range in $[0, 1]$, for any $x \in A$.
(3) There are only finitely many $\phi_j$, so for any $x \in A$, there is an open set $V$ of $x$ such that all but finitely many of $\phi_j$ are 0 on $V$.
(4) The sum of the functions at any point is 1.
(5) For each $\phi_j$, there is an element of $\mathcal{O}$ on which $\phi_j$ is supported.

It seems like we are done. However, we need one more step: we need to ensure that each of our functions has compact support. We do this by multiplying by the flat-top mountain function on $A$? **ASK ABOUT THIS**

## 7. Term Test 2 Rejects

## 8. Problem 1

Let the collection of points for which $f \neq g$ be denoted as $\{x_1, \ ..., \ x_n\}$. Let $m = g(x_1) + \cdots + g(x_n)$.

---

## 9. Problem 2

9.1. **Part A.** Pick some $\epsilon > 0$. Since $C$ is content-0, we can choose a finite collection of closed rectangles $R_k$ such that the sum of all $\text{vol}(R_k)$ is less than $\epsilon$. We claim that $\partial C$ is covered by the collection of $R_k$ as well. Let $R = \cup R_k$. Indeed, given some $c \in \partial C$, note that every neighbourhood of $c$ must intersect both $C$ and $C^C$, so $c \notin R^C$, as this is an open set that does not intersect $C$.

It follows by definition that $\partial C$ is also content-0.

9.2. **Part B.** Take the rationals in $[0, 1]$, in the context of $\mathbb{R}$. This set is countable, so it is measure-0, but its boundary is clearly $[0, 1]$, which is not measure-0.

---

## 10. Problem 3

Since $f$ is integrable, it is continuous except on a set of measure-0, which we denote $S$. Since $g$ is continuous, it follows that the composite $g \circ f$ is continuous for all $x \notin S$. Thus, $f \circ g$ is continuous except on a measure-0 set, so it is integrable.

---

## 11. Problem 4

Since $A$ is Jordan-measureable, the indicator function $\chi_A$ is integrable on some rectangle $K$ containing $A$. Thus, by the Reimann criteria for integration, we can pick some partition $P$ of $K$ such that:

$$U(\chi_A, P) - L(\chi_A, P) = \sum_{S \in P} M_S(\chi_A)\text{vol}(S) - \sum_{S \in P} m_S(\chi_A)\text{vol}(S) < \frac{1}{257}$$

Clearly, for some $S \in P$ such that $S \subset A$, we will have $M_S(\chi_A) = m_s(\chi_A) = 1$, and similarly, for $S \subset A^C$, we will have both $M_S$ and $m_S$ equal to 0. Finally, for the remaining $S$, which intersect both $A$ and $A^C$, we have $M_S = 1$ and $m_S = 0$, so:

$$\sum_{S \in P} M_S(\chi_A)\text{vol}(S) - \sum_{S \in P} m_S(\chi_A)\text{vol}(S) = \sum_{S \in P'} \text{vol}(S) < \frac{1}{257}$$

where $P'$ is the set of all rectangles intersecting both $A$ and $A^C$. Thus, taking $R = P$ and $R' = P'$, the proof is complete.

---

## 12. Problem 5

12.1. **Part A.** Since $\chi_B$ is integrable, it follows that $\partial B$ is measure-0, so the closed set $\overline{B} = B \cup \partial B$ containing $B$ is measure-0. Since this set is also bounded, it follows that it is compact, so it is content-0.

Thus, for any $\epsilon > 0$, we can pick some finite collection $S$ of rectangles covering $\overline{B}$, with sum of volumes less than $\epsilon$. We can take intersections of these rectangles, and extend the resulting set to a partition $P$ of a rectangle $R$ containing $\overline{B}$.

Clearly, the upper sum of $\chi_B$ on $R$ will be the sums of the volumes of rectangles intersecting $B$, which is precisely the subset of $P$ of rectangles obtained from intersecting elements of $S$. Clearly, the sum of volumes of these rectangles will be less than $\epsilon$. Since $\mathrm{vol}(B) \le U(\chi_B, P)$, for all $P$, it follows that $\mathrm{vol}(B)$ must equal 0, as we can make $U(\chi_B, P)$ arbitrarily small.

12.2. **Part B.** Let $A$ be the set of rationals in $[0, 1]$. Checking that $\chi_A$ satisfies the criteria is a simple exercise.

---

## 13. Problem 6

Suppose is not equal to 0. Then there exists some point $a$ at which $f(a) \neq 0$. Since $f$ is continuous, there is a neighbourhood of this point on which $f > 0$. Inside this neighbourhood, we can pick a rectangle $R$.

Thus, picking a partition containing $R$, we get a lower sum that is greater than 0, implying the integral itself must be greater than 0, which is a contradiction. Thus, we must have $f = 0$.

---

## 14. Problem 8

---

Since $A$ and $B$ are Jordan-measureable, it follows that $\chi_A$ and $\chi B$ are integrable. In addition, we have assumed that the functions $f_t(x) = \chi_{A_t}(x)$ and $g_t(x) = \chi_{B_t}(x)$, where $A_t$ and $B_t$ are the slices of $A$ and $B$, are integrable as well.

Recall that for any $t$, we have:

$$\int \chi_{A_t}(x) = \int \chi_{B_t}(x)$$

by assumption. Clearly, for some $t$, we have $\chi_{S_t}(x) = \chi_S(x, t)$. Thus, by Fubini's theorem:

$$\mathrm{vol}(A) = \int \chi_A = \int_{\mathbb{R}} \int_{R_A} \chi_A(x, t)\ dx\ dt = \int_{\mathbb{R}} \int_{R_B} \chi_B(x, t)\ dx\ dt = \mathrm{vol}(B)$$

and we are done. **Note:** I'm being quite sloppy with my notation in a few places, but the idea should be clear.

---

## 15. Problem 11

Let $E$ be the ellipsoid in question. This is a standard change of vbariables. Let:

$$f(x, y, z) = \left( \frac{x}{\sqrt{2}}, \frac{y}{\sqrt{3}}, \frac{z}{\sqrt{5}} \right)$$

Obviously, this function is bijective, and differentiable, with its differential having a non-zero determinant:

$$Df(x, y, z) = \mathrm{diag}\left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{5}} \right)$$

Finally, it obvious that:

$$f(C) = f(\{(x, y, z) \mid x^2 + y^2 + z^2 \le \}) = \left\{ \left( \frac{x}{\sqrt{2}}, \frac{y}{\sqrt{3}}, \frac{z}{\sqrt{5}} \right) \;\middle|\; x^2 + y^2 + z^2 \le 1 \right\} = E$$

Thus, by change of variables:

$$\text{vol}(E) = \text{vol}(f(C)) = \int_{f(C)} 1 = \int_C |\det Df|$$

From above, $|\det Df| = \frac{1}{\sqrt{30}}$. Thus:

$$\text{vol}(E) = \frac{1}{\sqrt{30}} \int_C 1$$

Since $C$ is simply a sphere with radius 1, we know $\int_C 1 = \frac{4}{3}\pi$ (we could also show this with a spherical coordinate transform, but we are lazy). Thus, the volume of the ellipsoid is $\frac{1}{\sqrt{30}}\frac{4\pi}{3}$.

---

## 16. Problem 12

16.1. **Part A.** This is an immediate consequence of the fundamental theorem of calculus. Let $g_1(x,y) = \partial_x \partial_y f(x,y)$, and let $g_2(x,y) = \partial_y \partial_x f(x,y)$. We know both of these functions are continuous. Hence, by Fubini's theorem:

$$\int_R g_1(x,y) = \int_{[c,d]} \int_{[a,b]} \partial_x \partial_y f(x,y)\, dx\, dy = \int_{[c,d]} (f(b,y) - f(a,y))\ dy = f(b,d) - f(a,d) - f(b,c) + f(a,c)$$

An almost identical calculation shows that $\int_R g_2$ yields the same result. Thus, we have the desired equality.

16.2. **Part B.** Suppose there is some $(a,b)$ at which $\partial_x \partial_y f - \partial_y \partial_x f > 0$. Since this function is continuous, there must be a neighbourhood around $(a,b)$ on which it is positive. Taking the integral on a rectangle contained in this neighbourhood gives some positive number, but this contradicts Part A. Thus, not $(a,b)$ exists.

Identical logic shows that an $(a,b)$ at which the difference is negative cannot exist. Thus, $\partial_x \partial_y f = \partial_y \partial_x f$ for all points.

---

## 17. Problem 13

Let $f(x,y) = \sqrt{x^2+y^2}$. Note that $f \circ T_\theta = f$:

$$(f \circ T_\theta)(x,y) = f(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta) = \sqrt{(x^2+y^2)(\sin^2\theta + \cos^2\theta)} = \sqrt{x^2+y^2} = f(x,y)$$

Since $B$ is Jordan-measureable, it bounded, so it is contained in some rectangle $R$. Since $f$ is continuous, it has a maximum value on $R$. Since $f$ is the radial distance of a point $(x,y)$ from the origin, it therefore follows there is some circle $C$ containing $R$, and thus $B$. It is easy to see that $T_\theta(C) = C$.

Now, all that is left to do is a change of variables. Clearly, $T_\theta$ is a diffeomorphism, and has determinant 1. Thus:

$$\text{vol}(B) = \int_C \chi_B = \int_{T_\theta(C)} \chi_B = \int_C \chi_B \circ T_\theta = \int_C \chi_{T_\theta(B)} = \text{vol}(T_\theta B)$$

---

## 18. Problem 14

This is a straightforward application of definitions.

---

## 19. Problem 15

Easy

---

## 20. Problem 16

Easy

---

## 21. Problem 17

Easy

---