MATH 395 Notes

Faye Jackson

August 31, 2020

1 Metric Spaces

1.1 Definition

Definition. *A set X is called a metric space provided that it is equipped with a function $d : X \times X \to [0, \infty)$ such that*

1. *For all $p, q \in X$ we have $d(p, q) = 0$ if and only if $p = q$*
2. *$d(p, q) = d(q, p)$ for all $p, q \in X$.*
3. *For all $p, q, r \in X$ we have*

$$d(p, q) \leq d(p, r) + d(r, q)$$

We call d the metric on X. Formally we might write that (X, d) is a metric space, since a set X may admit many different metrics on it.

Example. Let $X = \mathbb{R}^n$ for some $n \in \mathbb{N}$. If $p = (p_1, \ldots, p_n)$ and $q = (q_1, \ldots, q_n)$ then we define:

$$d_2(p, q) = \left[\sum_{j=1}^{n} (q_j - p_j)^2\right]^{\frac{1}{2}} = \|p - q\| = \langle q - p, q - p\rangle^{\frac{1}{2}}$$

This is commonly called the ℓ^2 metric on $\mathbb{R}^n$. The triangle inequality follows from Cauchy-Schwartz. Setting $x = p - r$ and $y = r - q$, then $x = y = p - q$ and we also have:

$$\|x+y\|^2 \leq (\|x\| + \|x\|)^2$$
$$\|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|$$

But since we know from Cauchy-Schwarz that $\langle x, y\rangle \leq \|x\|\|y\|$, so we win!

We can put another metric on R^n, namely the ℓ^s metric for any $1 \leq s < \infty$:

$$d_s(p, q) = \left[\sum_{j=1}^{n} |q_j - p_j|^s\right]^{\frac{1}{s}}$$

This is called the ℓ^s metric. There is also the ℓ^∞ metric denoted as:

$$d_\infty(p, q) = \max_{1 \leq j \leq n} |q_j - p_j|$$

1.2 Topology on metric spaces

Definition. *A topology on a set X is some collection of subsets $\mathscr{T} \subseteq P(X)$, which we will call the open subsets of X, such that:*

- *$\emptyset$ and X are both open.*
- *Given any arbitrary family of open sets $\{U_i\}_{i\in I}$, their union $\bigcup_{i\in I} U_i$ is an open set*
- *Given any finite collection of open sets, $U_1, \ldots, U_n$, then their intersection $\bigcap_{i=1}^{n} U_i$ is open.*

Definition. *Let (X, d) be a metric space. We define a topology on X as follows:*

- *For $x_0 \in X$ and $\varepsilon > 0$ we define the ε-neighborhood of x_0 as:*

$$N_\varepsilon(x_0) := \{x \in X \mid d(x, x_0) < \varepsilon\}$$

- *A subset $U \subseteq X$ is called open provided that for every $p \in U$ there exists some $\varepsilon > 0$ so that $N_\varepsilon(p) \subseteq U$.*

Proof that this is a topology. The first property follows nearly trivially.

- Fix some arbitrary family of open sets $\{U_i\}_{i\in I}$. Fix some $p \in \bigcup_{i\in I} U_i$, then there exists some $j \in I$ so that $p \in U_j$. Since U_j is open there exists some $\varepsilon > 0$ so that:

$$N_\varepsilon(p) \subseteq U_j \subseteq \bigcup_{i\in I} U_i$$

  And so we are done ☺

- Let $p \in \bigcap_{i=1}^n U_i$ for some finite collection of open sets $U_1, \ldots, U_n$. Then $p \in U_j$ for all $1 \leq j \leq n$, and so there exists an $r_j > 0$ for each j such that:

$$N_{r_j}(p) \subseteq U_j$$

  Take $r = \min(r_1, \ldots, r_n$. Then for all j we have $N_r(p) \subseteq N_{r_j}(p) \subseteq U_j$. And so:

$$N_r(p) \subseteq \bigcap_{i=1}^n U_i$$

  just as desired.

Remark. This third property is not true for infinite collections! What part of the proof breaks and provide a counter-example.

With this we are done.

Exercise. *Also, as an exercise, show that for any $r > 0$ we have $N_r(p)$ is open.*

Definition. *We say a subset $C \subseteq X$ of a topological space is closed provided that its complement $X \setminus C$ is open.*

Remark. By Demorgan's laws we get three properties of closed sets:

- $\emptyset$ and X are both closed
- If $\{C_i\}_{i\in I}$ is a collection of closed sets then $\bigcap_{i\in I} C_i$ is closed
- If $C_1, \ldots, C_n$ is a finite collection of closed sets then $\bigcup_{i=1}^n C_i$ is a closed set.

The proof is left as an exercise ☺

1.3 Limit Points / Accumulation Points

Definition. *A point p is called a limit point of a set E provided that every neighborhood of p contains a point $q \neq p$ such that $q \in E$.*

Example. Let $E = [0, 1) \cup \{2\}$. Then 1 is a limit point of E (note that $1 \notin E$), and also 2 is *not* a limit point of E even though $2 \in E$.

Definition. *When $p \in E$ is not a limit point of E, p is called an isolated point of E.*

Definition. *An interior point of E is a point $p \in E$ such that there exists $r > 0$ so that $N_r(p) \subseteq E$. Thus a set is open exactly when all its points are interior points. The set of all interior points of a set E is often denoted by $\mathring{E}$, this is called the interior of E.*

Example. This depends on the entire metric space

- Let $E = [0, 1) \cup \{2\}$ and $X = [0, \infty)$. Then 0 is an interior point of E (since $N_r(0) = [0, r) \subseteq E$ is r is small enough). Thus $\mathring{E} = [0, 1)$.
- Let $E = [0, 1) \cup \{2\}$ and $X = \mathbb{R}$. Then 0 is not an interior point of E, since any neighborhood of 0 will contain negative numbers, which are not contained in E.

Thus we conclude that the notion of interior (open or closed) depends on the ambient space.

Definition. *A set E is bounded provided that there eixsts a point $x \in X$ and a number $M > 0$ such that $E \subseteq N_M(x)$.*

Definition. *A set $E \subseteq X$ is dense provided that every point of X is either a limit point of E or an element in E.*

Example. Let $X = [0, 1) \cup \{\pi\}$ then $X \cap \mathbb{Q} \cup \{\pi\}$ is dense in X. Notice that $X \cap \mathbb{Q}$ is not dense in X.

Theorem. *If p is a limit point of a set E, then every neighborhood of p contains infinitely many points of E.*

Exercise. *Prove this*

Corrolary. *A finite set can have no limit points*

Theorem. *A set E is closed if and only if every limit point of E is contained in E.*

Proof. Let's do it! We will use X as our ambient space.

($\Rightarrow$) Let E be closed and suppose p is a limit point of E. If $p \notin E$ then $p \in X \setminus E$, which is open, and so there exists an $r > 0$ such that $N_r(p) \subseteq X \setminus E$. Therefore $N_r(p) \cap E = \emptyset$, but this contradicts the fact that p is a limit point. Therefore $p \in E$ as desired.

($\Leftarrow$) Suppose that every limit point belongs to E and take $p \in X \setminus E$. Since p is not a limit point of E there must exist some $r > 0$ such that $N_r(p) \cap E = \emptyset$. But then $N_r(p) \subseteq X \setminus E$. Therefore $X \setminus E$ is open, and E is closed.

Awesome! We win ☺

Definition. *A set E is called perfect if E is closed and every point of E is a limit point. In other words, E consists exactly of its limit points.*

Example. $[0, 1]$ is perfect in $\mathbb{R}$, but $[0, 1] \cup \{\pi\}$ is not.

Example. Let $X = \mathbb{R}^2 = \mathbb{C}$. Consider the following sets

a) The set of all complex numbers $|z| < 1$

b) The set of all complex numbers $|z| \leq 1$

c) A finite set $F \subseteq \mathbb{C}$

d) The set of all integers $\{(n, 0) \mid n \in \mathbb{N}\}$

e) The set $z_n = \frac{1}{n}$ where $n \in \mathbb{N}$

f) The set of all complex numbers

g) The line segment (a, b) for $a, b \in \mathbb{R}$. That is the set of points $z \in \mathbb{C}$ such that $\mathrm{Im}(z) = 0$ and $a < \mathrm{Re}(z) < b$

	Closed	Open	Bounded	Perfect
a)	✗	✓	✓	✗
b)	✓	✗	✓	✓
c)	✓	✗	✓	✗
d)	✓	✗	✗	✗
e)	✗	✗	✓	✗
f)	✓	✓	✗	✓
g)	✗	✗	✓	✗

MATH 395 Notes

Faye Jackson

September 2, 2020

1 Introduction

- Office Hours
  - Monday 8-9am
  - Wednesday 4-5pm
  - Beginning of Friday lecture
- First HW will be posted on Friday
- TAs are still not decided on
- Further info on the waitlist to come

2 Continuing Metric Spaces

2.1 Last Time

We defined metrics $d : X \times X \to [0, \infty)$ with three special properties, and we saw that this gave us a topology on X.

- Open sets, given $p \in \mathcal{O}$ we have some $\delta > 0$ so that $N_\delta(p) \subseteq \mathcal{O}$.
- Closed sets are the complements of open sets
- Limit points, p is a limit point of E if every δ-neighborhood of p intersects E in a point $q \neq p$
- Closed sets are exactly the sets where every limit point belongs to the set.

2.2 Closures!

Definition. *If X is a metric space and $E \subseteq X$ we denote by E' the set of limit points of X. The closure of E is the set $\overline{E} = E \cup E'$.*

Example. Here are some examples to look at!

- Let $E = (0, 1] \subseteq \mathbb{R}$ then $E' = [0, 1] = \overline{E}$
- Let $E = (0, 1] \cup \{2\} \subseteq \mathbb{R}$. Then $E' = [0, 1]$ and $\overline{E} = [0, 1] \cup \{2\}$.

Theorem. *Let X be a metric space and $E \subseteq X$. Then:*

a) $\overline{E}$ is closed

b) $E = \overline{E}$ if and only if E is closed.

c) If $E \subseteq F$ and F is closed then $\overline{E} \subseteq F$.

Proof. Let's go!

a) Let $q \in \overline{E}^c$. Then $q \notin E' \cup E$. Thus there exists a $\delta > 0$ so that $N_\delta(q) \cap E = \emptyset$. Since $N_\delta(q)$ is open we also know that $N_\delta(q) \cap E' = \emptyset$. Therefore $N_\delta(q) \cap \overline{E} = \emptyset$ and so $\left(\overline{E}\right)^c$ is open as desired.

b) Easy exercise

c) If $E \subseteq F$ and F is closed, then $E' \subseteq F$ because any limit point of E is also a limit point of F. Therefore $\overline{E} \subseteq F$.

Theorem. *Let E be a nonempty set of real numbers which is bounded above. Then $y = \sup E$ is in $\overline{E}$. Hence $y \in E$ if E is closed.*

Proof. If $y \in E$ then we are done because $E \subseteq \overline{E}$. If $y \notin E$ then for any $\varepsilon > 0$ there exists some $x \in E$ so that:

$$y - \varepsilon < x < y$$

But this means that $x \in N_\varepsilon(y)$, and so $N_\varepsilon(y) \cap E \neq \emptyset$. This implies that $y \in E'$, and so we are done since $E' \subseteq \overline{E}$.

2.3 Compact subsets of metric spaces

Definition. *We need a couple definitions!*

- *An open cover of a set E in a metric space X is a collection $\{G_\alpha\}_{\alpha \in A}$ of open sets such that:*

$$E \subseteqq \bigcup_{\alpha \in A} G_\alpha$$

- *A subset $E \subseteq X$ is called compact provided that every open cover of E admits a finite subcover. That is we can find a finite subcollection $\{G_{\alpha_i}\}_{1 \le i \le n}$ of $\{G_\alpha\}_{\alpha \in A}$ such that $\{G_{\alpha_i}\}_{1 \le i \le n}$ covers E.*

Theorem. *Compact subsets of metric spaces are closed and bounded*

Proof of Closed. Let $K \subseteq X$ be compact and let $q \in K^c$. For each $p \in K$ there exists two subsets U_p and W_p such that $p \in U_p$, $q \in W_p$ and $U_p \cap W_p = \emptyset$. Here we use that metric spaces are Hausdorff. We can concretely take $U_p = N_\delta(p)$ and $W_p = N_\delta(q)$ with $\delta < \frac{1}{2}d(p,q)$.

Then in fact $\{U_p\}_{p \in K}$ is an open cover of K. By compactness there exists a finite subcover $U_{p_1}, \dots U_{p_n}$ that covers K. Then let:

$$W = \bigcap_{i=1}^{n} W_{p_i}$$

Then this W is open and $W \cap U_{p_j} = \emptyset$ for all $1 \le j \le n$. Thus we must have $W \cap K = \emptyset$, meaning that $W \subseteq K^c$ and K^c is open.

Proof of Boundedness. Let $x \in X$ be arbitrary. The family of sets $\{N_n(x)\}_{n \in \mathbb{N}}$ is an open cover of E since $\mathbb{N}$ is unbounded. Thus by compactness E has a finite subcover, and so $E \subseteq N_k(x)$ for some $k \in \mathbb{N}$.

The main question for the rest of this section: Is the converse true? If not, what should be a workable criterion for compactness in metric spaces?

In fact it is true on $\mathbb{R}^n$ by Heine-Borel. But not the converse, particularly in infinite dimensions!

Theorem. *Closed subsets of compact sets are compact.*

Proof. Let $C \subseteq K$ be a closed subset of a compact set K and let $\{G_\alpha\}_{\alpha \in A}$ be an open cover of C. Then $\{G_\alpha\}_{\alpha \in A} \cup C^c$ is an open cover for K. Thus by compactness of K there exists $\alpha_1, \ldots, \alpha_n$ such that:

$$K \subseteq C^c \cup \bigcup_{i=1}^{n} G_{\alpha_i}$$
$$C \subseteq \bigcup_{i=1}^{n} G_{\alpha_i}$$

Therefore C is comapact.

Theorem (Finite intersection property)**.** *If $\{K_\alpha\}_{\alpha \in A}$ is a collection of compact sets such that the intersection of any finite subcollection of $\{K_\alpha\}_{\alpha \in A}$ is nonempty. Then, the intersection $\bigcap_{\alpha \in A} K_\alpha$ is nonempty*

Example. If $E_n = \left(0, \frac{1}{n}\right]$ then E_n has the finite intersection property since they are nested and each of them are nonempty. But $\bigcap_{n \in \mathbb{N}} E_n = \emptyset$.

Proof. Suppose that $\bigcap_{\alpha \in A} K_\alpha = \emptyset$. Then $\bigcup_{\alpha \in A} K_\alpha^c = X$, and so $\{K_\alpha^c\}_{\alpha \in A}$ is an open cover for $K_{\alpha_\star}$ where $\alpha_\star \in A$ is arbitrary. This holds because compact subsets of metric spaces are closed.

By compactness of $K_{\alpha_\star}$ there exists some $\alpha_1, \ldots, \alpha_n$ such that:

$$K_{\alpha_\star} \subseteq \bigcup_{i=1}^{n} K_{\alpha_i}^c$$

Thus the finite intersection:

$$K_{\alpha_\star} \cap \bigcap_{i=1}^{n} K_{\alpha_i} = \emptyset$$

This contradicts the finite intersection property. Oops! We win.

Theorem 1 (Compactness $\implies$ sequential compactness)**.** *Let K be a compact set and let $\{x_n\}_{n \in \mathbb{N}}$ be a sequence of points in K. Then there exists a convergent subsequence $\{x_{n_k}\}_{k \in \mathbb{N}}$ of $\{x_n\}_{n \in \mathbb{N}}$ that converges to a point in K.*

Proof. Suppose that $\{x_n\}$ has no limit point in K. This means that for any $p \in K$, there exists some δ_p such that $N_{\delta_p}(p)$ contains at most one point of the sequence $\{x_n\}$. The collection $\{N_{\delta_p}(p)\}_{p \in K}$ is an open cover of K.

By compactness we have some $p_1, \ldots, p_n$ such that:

$$K \subseteq \bigcup_{i=1}^{n} N_{\delta_{p_i}}(p_i)$$

But this must mean that K contains at most n points of the sequence $\{x_n\}$. This means that $\{x_n\}$ takes at most n values. Thus x_n must take one value infinitely many times, and so x_n has a convergent subsequence.

On the other hand if $\{x_n\}$ has a limit point $p \in K$, then for every $k \in \mathbb{N}$ there exists some x_{n_k} such that $d(x_{n_k}, p) < \frac{1}{k}$. Clearly $\{x_{n_k}\}$ is a convergent subsequence and so we win.

Remark. Is the converse true? Yes! But only in metric spaces.

3 Compactness in $\mathbb{R}^n$

Theorem 2 (Nested interval property on $\mathbb{R}$)**.** *Suppose that $I_n = [a_n, b_n]$ is a nested sequence of closed intervals, that is $I_n \supseteq I_{n+1}$. Then $\bigcap_{n=1}^{\infty} I_n$ is nonempty*

Proof. We know $\{a_n\}$ is an increasing sequence thta is bounded by b_j. Let $x = \sup_{n \in \mathbb{N}} a_n$. Then $a_n \leq x$ for all n.

Also $\{b_n\}$ is decreasing so $a_n \leq b_n \leq b_m$ for all $n \geq m$. Taking the supremum in n we get $x \leq b_m$ for all m. Therefore $a_n \leq x \leq b_n$ for all $n \in \mathbb{N}$, giving us that:

$$x \in \bigcap_{i=1}^{\infty} I_n$$

Handout 1

- **What is a topology on a set X?** Let X be a set. A topology on X is a collection $\mathcal{T}$ of subsets of X that are called *open sets* satisfying the following three conditions:

  C1) $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$,

  C2) Given a collection $O_\alpha \in \mathcal{T}$ of index sets, then $\cup_\alpha O_\alpha \in \mathcal{T}$ as well; We say that $\mathcal{T}$ is closed under unions,

  C3) Given a *finite* collection of open set $O_1, \ldots, O_n$, then $\cap_1^n O_n \in \mathcal{T}$; We say that $\mathcal{T}$ is closed under finite intersections.

- A topology can be equivalently defined by specifying the collection of *closed sets* which satisfy the same conditions as above except that we switch unions $\cup$ with intersections $\cap$ in conditions C2) and C3). The couple $(X, \mathcal{T})$ is called a topological space, or sometimes we just say X is a topological space if we're only playing with one agreed upon topology

- A space X can have more than one topology defined on it. A topology $\mathcal{T}_1$ is said to be finer or stronger than $\mathcal{T}_2$ if $\mathcal{T}_2 \subset \mathcal{T}_1$ (we say $\mathcal{T}_2$ is courser or weaker). Notice that the trivial topology $\{\emptyset, X\}$ is the weakest topology on X.

- One way to describe a topology on a set X is to define precisely all open sets. This is what we did for metric spaces. Occasionally, we want to define the smallest topology that designates a particular collection $\mathcal{B}$ of subsets of X as open. This is done as follows:

Q1) Let $\overline{\mathcal{B}}$ be the collection of subsets of X that contains the empty set, X, as well as all sets obtained as finite intersections of elements of $\mathcal{B}$. Show that the collection $\mathcal{T}$ obtained by taking unions of elements of $\overline{\mathcal{B}}$ is a topology on X.

Q2) Show that any other topology on X that contains $\mathcal{B}$ as open sets, contains $\mathcal{T}$. We call $\mathcal{T}$ the topology generated by $\mathcal{B}$. It is the coarsest topology containing $\mathcal{B}$.

- (Product Topology) One example where this construction is useful is to define a topology on the product of topological spaces. Suppose $(X_\alpha, \mathbb{T}_\alpha)$ are topological spaces for $\alpha \in A$ (where A is an index set that could be infinite). We would like to define a "natural" topology on $\prod_\alpha X_\alpha$. One reasonable requirement is that the *cylindrical sets* are open (cylindrical sets are those of the form $\prod_\alpha U_\alpha$ where all the U_α are open in X_α and all but one of them is equal to X_α. The topology generated by this collection is called the product or Tychonoff topology.

Q3) Consider the product topology on $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ as defined above. Why is this the same as the standard topology on $\mathbb{R}^2$ defined in class.

- We saw in class that the interval $[0, 1)$ is not open in $\mathbb{R}$, but is open relative to the half-line $[0, \infty)$ (taking the usual metric on $\mathbb{R}$ and $[0, \infty)]$).
Let us try to formalize and generalize this.

  Let (X, d) be a metric space and $Y \subset X$. Y is a metric space itself, by restricting the metric d to $Y \times Y$.

Q4) Let $E \subset Y$. We say that E is open relative to Y if it is open in the metric space (Y, d). Untangle what this definition means in terms of $N_\delta(p)$ neighborhood of a point $p \in E$. Deduce that if there is an open subset G of X, then $G \cap Y$ is open relative to Y.

Q5) Show that E is open relative to Y if and only if there exists an open subset G of X such that $E = G \cap Y$.

Q6) Compactness on the other hand behaves better. Suppose that $K \subset Y \subset X$. Then K is compact relative to X if and only if it is compact relative to Y.

Remark: As such, we always need to specify the ambient space when we talk about open/closed sets (that's why we always say "E is an open subset of X"), but we can make statements like "K is compact (or a compact metric space)" without the need to specify the ambient space.

MATH 395 Notes

Faye Jackson

September 4, 2020

Exercise 1. *Prove Q1*

Proof. Let's go!

- Note that $\emptyset \in \overline{\mathfrak{B}}$ is in particular an element of the set $\mathcal{T}$. Likewise $X \in \mathcal{T}$

- Consider any collection $\{U_\alpha\}_{\alpha \in A}$ where each U_α is an element of $\mathcal{T}$. Then for each α there are basis sets $\{\overline{B}_i\}_{i \in I_\alpha} \subseteq \overline{\mathfrak{B}}$ so that:

$$U_\alpha = \bigcup_{i \in I_\alpha} \overline{B}_i$$

  Therefore we have that:

$$\bigcup_{\alpha \in A} U_\alpha = \bigcup_{\alpha \in A} \bigcup_{i \in I_\alpha} \overline{B}_i = \bigcup_{i \in \bigcup_{\alpha \in A} I_\alpha} \overline{B}_i$$

  And therefore by definition of $\mathcal{T}$ we know the union of the $\{U_\alpha\}$ is an element of $\mathcal{T}$.

- Consider any finite collection $U_1, \ldots, U_n$ in $\mathcal{T}$. For each $1 \leq i \leq n$ there are basis sets $\{\overline{B}_\alpha\}_{\alpha \in A_i}$ each in $\overline{\mathfrak{B}}$. If any of the $\overline{B}_\alpha$ for $\alpha \in A_i$ are the empty set then they don't effect U_i, and if any of them are the whole space then that $U_i = X$ and it doesn't effect the whole intersection.

Thus we can assume that there exists $\{B_j\}_{1 \leq j \leq m_\alpha}$ in $\mathfrak{B}$ such that:

$$\begin{aligned}
\overline{B}_\alpha &= \bigcap_{j=1}^{m_\alpha} B_j \\
U_i &= \bigcup_{\alpha \in A_i} \overline{B}_\alpha \\
&= \bigcup_{\alpha \in A_i} \bigcap_{j=1}^{m_\alpha} B_j
\end{aligned}$$

Therefore we can write by Demorgan:

$$\begin{aligned}
\bigcap_{i=1}^{n} U_i &= \bigcap_{i=1}^{n} \bigcup_{\alpha \in A_i} \bigcap_{j=1}^{m_\alpha} B_j \\
&= \bigcup_{(\alpha_1, \dots, \alpha_n) \in \prod_{i=1}^{n} A_i} \bigcap_{i=1}^{n} \bigcap_{j=1}^{m_{\alpha_i}} B_j
\end{aligned}$$

And since the finite intersection of finite intersections is a finite intersection we win, this is open.

Exercise 2. *Show Q2*

Proof. Fix a topology $\mathbb{T}$ on X which contains each element of $\mathfrak{B}$. Fix some open set $U \in \mathcal{T}$. Then we know there is some collection $\{\overline{B}_\alpha\}_{\alpha \in A}$ each in $\overline{\mathfrak{B}}$ such that:

$$U = \bigcup_{\alpha \in A} \overline{B}_\alpha$$

Thus we merely just need to show that $\overline{\mathfrak{B}} \subseteq \mathbb{T}$ since $\mathbb{T}$ is closed under arbitrary unions:

- We know that $\emptyset$ and X are elements of $\mathbb{T}$ since $\mathbb{T}$ is a topology
- In the other case for $\overline{B} \in \overline{\mathfrak{B}}$ we have that for some $B_1, \dots, B_n$ in $\mathfrak{B}$ that:

$$\overline{B} = \bigcap_{i=1}^{n} B_i$$

Since $\mathbb{T}$ contains each B_i and it is closed under finite intersection we then know that $\overline{B}$ is in $\mathbb{T}$ as desired.

Thus we win! We have that $\mathbb{T} \subseteq \mathcal{T}$.

Exercise 3. *Show Q3. That is show the product topology on $\mathbb{R}^2$ agrees with the Euclidean topology on $\mathbb{R}^2$.*

Proof. Call the product topology $\mathcal{T}_\pi$ and the Euclidean topology $\mathcal{T}_\mathcal{E}$. We proceed by two-way containment.

($\subseteq$) We know by $Q2$ that to show $\mathcal{T}_\pi \subseteq \mathcal{T}_\mathcal{E}$ it suffices to show that each cylindrical set is an open set in the Euclidean topology. There are two cases:

- Suppose that U is open in $\mathbb{R}$. We must show that $U \times \mathbb{R}$ is open in $\mathbb{R}^2$ with the Euclidean topology. Fix $(x, y) \in U \times \mathbb{R}$. Then $x \in U$, so there exists some $\varepsilon > 0$ so that $N_\varepsilon(x) \subseteq U$. We claim that $N_\varepsilon(x, y) \subseteq U \times R$. Fix $(v, w) \in N_\varepsilon(x, y)$. Then we know that:

$$\begin{aligned} d(x, v) &= |x - v| = \sqrt{(x-v)^2} \\ &\leq \sqrt{(x-v)^2 + (y-w)^2} = d((x, y), (v, w)) < \varepsilon \end{aligned}$$

  Therefore $v \in N_\varepsilon(x) \subseteq U$. Since $v \in U$ and $w \in \mathbb{R}$ we know that $(v, w) \in U \times \mathbb{R}$ as desired.

- Suppose that U is open in $\mathbb{R}$. We must show that $\mathbb{R} \times U$ is open in $\mathbb{R}^2$ with the Euclidean topology. Fix $(x, y) \in \mathbb{R} \times \mathbb{R}$. Then $y \in U$, so there exists some $\varepsilon > 0$ so that $N_\varepsilon(y) \subseteq U$. We claim that $N_\varepsilon(x, y) \subseteq U \times R$. Fix $(v, w) \in N_\varepsilon(x, y)$. Then we know that:

$$\begin{aligned} d(y, w) &= |y - w| = \sqrt{(y-w)^2} \\ &\leq \sqrt{(x-v)^2 + (y-w)^2} = d((x, y), (v, w)) < \varepsilon \end{aligned}$$

  Therefore $w \in N_\varepsilon(y) \subseteq U$. Since $w \in U$ and $v \in \mathbb{R}$ we know that $(v, w) \in \mathbb{R} \times U$ as desired.

($\supseteq$) Fix some open set $U \subseteq \mathbb{R}^2$ with the Euclidean topology. Fix some $(x, y) \in U$. Then there is an $\varepsilon > 0$ so that $N_\varepsilon(x, y) \subseteq U$. Then set $\delta := \frac{\varepsilon}{\sqrt{2}}$. Consider then this open set in the product topology:

$$V_{(x,y)} = (N_\delta(x) \times \mathbb{R}) \cap (\mathbb{R} \times N_\delta(y)) = N_\delta(x) \times N_\delta(y)$$

It is clear that $(x, y) \in V_{(x,y)}$. Now take $(a, b) \in V_{(x,y)}$. We then know that $|a - x| < \frac{\varepsilon}{\sqrt{2}}$ and $|y - b| < \frac{\varepsilon}{\sqrt{2}}$. We then must have the following:

$$\begin{aligned} (a - x)^2 &< \frac{\varepsilon^2}{2} \\ (b - y)^2 &< \frac{\varepsilon^2}{2} \\ (a - x)^2 + (b - y)^2 &< \varepsilon^2 \\ d((a, b), (x, y)) &< \varepsilon \end{aligned}$$

Therefore $(a, b) \in N_\varepsilon(x, y) \subseteq U$. This shows that $V_{(x,y)} \subseteq U$. This lets us write that:

$$U = \bigcup_{(x,y) \in U} V_{(x,y)}$$

Thus since $\mathcal{T}_\pi$ is a topology and each $V_{(x,y)}$ is open in $\mathcal{T}_\pi$ we win! We have that U is open in $\mathcal{T}_\pi$.

With this we win!

MATH 395 Notes

Faye Jackson

September 9, 2020

3 Compactness on $\mathbb{R}^d$

Last time we proved the nested interval property on $\mathbb{R}$, namely

Theorem (Nested Interval Property). *Let $I_n = [a_n, b_n]$ be a sequence of closed and bounded intervals that is nested, aka $I_n \supseteq I_{n+1}$. Then we have that:*

$$\bigcap_{n=1}^{\infty} I_n \neq \emptyset$$

Note that we need "closedness." Take $I_n = \left(0, \frac{1}{n}\right]$. In fact what we really need is compactness.

Definition. *A closed box in $\mathbb{R}^d$ is a set of the form:*

$$\prod_{j=1}^{d} [a_j, b_j]$$

Corrolary (The nested box property of $\mathbb{R}^d$). *Let B_n be a sequence of closed and nested boxes. Then:*

$$\bigcap_{n=1}^{\infty} \neq \emptyset$$

Great!

Proof. Let:

$$B_n = \prod_{j=1}^{d} \left[a_j^{(n)}, b_j^{(n)}\right]$$

$B_n \subseteq B_{n+1}$ implies for every $1 \leq j \leq d$ that the intervals $I_j^{(n)} = \left[a_j^{(n)}, b_j^{(n)}\right]$ are nested. By the previous theorem, for every $1 \leq j \leq d$ there exists some:

$$x_j \in \bigcap_{n=1}^{\infty} I_j^{(n)}$$

Therefore:

$$x = (x_1, \ldots, x_d) \in \bigcap_{n=1}^{\infty} B_n$$

And so we win!

Definition. *Define in a metric space for any subset E of a metric space X the diameter when the following supremum exists:*

$$\operatorname{diam} E = \sup_{x,y \in E} d(x, y)$$

Great!

Exercise. *Show that for any box $B = \prod_{j=1}^{d} [a_j, b_j]$ that:*

$$\operatorname{diam} B = \left[\sum_{j=1}^{d} (b_j - a_j)^2\right]^{\frac{1}{2}}$$

Where we use the standard Euclidean metric on $\mathbb{R}^d$

Proof. We will do this with induction on d

- Suppose $d = 1$. We wish to prove that $\operatorname{diam}[a, b] = |b - a| = b - a$. Note that $b - a$ is in the set we are taking a supremum over, and so we merely need to show it is an upper bound. Fix $x, y \in [a, b]$. Without loss of generality take $y \geq x$. Then note that:

$$b - a = (b - y) + (y - x) + (x - a) \geq y - x$$

And so we win

- Suppose that the result holds for $d \in \mathbb{N}$. We must show it holds for $d+1$. Note then that $a = (a_1, \ldots, a_{d+1})$ and $b = (b_1, \ldots, b_{d+1})$ are in B, and so:

$$d(a,b) = \left[\sum_{j=1}^{d+1} (b_j - a_j)^2\right]^{\frac{1}{2}}$$

  Is in the set we are taking a supremum over. We need only show that it is a maximum. Fix $x = (x_1, \ldots, x_{d+1})$ and $y = (y_1, \ldots, y_{d+1})$ in the box B and without loss of generality assume $y_{d+1} \geq x_{d+1}$.

  Define $x' = (x_1, \ldots, x_d)$ and $y' = (y_1, \ldots, y_d)$. Then we have:

$$d(x', y') \leq \delta := \operatorname{diam} \prod_{j=1}^{d} [a_j, b_j] = \left[\sum_{j=1}^{d} (b_j - a_j)^2\right]^{\frac{1}{2}}$$

  Now note that:

$$\begin{aligned} d(x,y) &= \sqrt{(d(x', y'))^2 + (y_{d+1} - x_{d+1})^2} \\ &\leq \sqrt{\delta^2 + (b_{d+1} - a_{d+1}^2} \\ &= \left(\left[\sum_{j=1}^{d} (b_j - a_j)^2\right] + (b_{d+1} - a_{d+1})^2\right)^{\frac{1}{2}} \\ &= \left[\sum_{j=1}^{d+1} (b_j - a_j)^2\right]^{\frac{1}{2}} = d(a,b) \end{aligned}$$

  But this is exactly what we want ☺

Awesome!

Theorem. *Every closed box in* $\mathbb{R}^d$ *is compact.*

Proof. Let $B = \prod_{j=1}^{d} [a_j, b_j]$ be any closed box. Set:

$$\delta_0 := \operatorname{diam} B = \left[\sum_{j=1}^{d} (b_j - a_j)^2 \right]^{\frac{1}{2}}$$

Suppose for the sake of contradiction that $\{G_\alpha\}_{\alpha \in A}$ is an open cover of B that has no finite subcover

Split B into 2^d subboxes of equal size. That is let $c_j = \frac{a_j + b_j}{2}$. Then the subboxes are $\prod_{j=1}^{d} I_j$ where $I_j \in \{[a_j, c_j], [c_j, b_j]\}$.

Since B cannot be covered by any finite collection of the $\{G_\alpha\}_{\alpha \in A}$, there must exist a subbox, B_1 such that B_1 cannot be covered by any finite subcollection of the $\{G_\alpha\}_{\alpha \in A}$. Note also that $\operatorname{diam} B_1 = \frac{\operatorname{diam} B}{2}$. Set $\delta_1 = \operatorname{diam} B_1$.

Continue inductively, having constructed $B \supseteq B_1 \supseteq B_2 \supseteq \cdots \supseteq B_n$ such that $\operatorname{diam} B_n = \delta_n = \frac{\operatorname{diam} B}{2^n}$ and B_n cannot be covered by any finite collection of the $\{G_\alpha\}_{\alpha \in A}$. We construct B_{n+1} by splitting B_n into 2^d subboxes of equal size as in the previous paragraph and noting that one of those subboxes cannot be covered by any finite collection of the $\{G_\alpha\}_{\alpha \in A}$. Let B_{n+1} be this subbox of B_n. Also note:

$$\operatorname{diam} B_{n+1} = \frac{\operatorname{diam} B_n}{2} = \frac{\operatorname{diam} B}{2^{n+1}}$$

This is a sequence of closed nested boxes. Applying the nested box property we know that $\bigcap_{n=1}^{\infty} B - n \neq \emptyset$.

Claim. *$\bigcap_{n=1}^{\infty} B_n$ is a singleton x.*

Proof. Suppose $x, y \in \bigcap_{n=1}^{\infty}$. Then $x, y \in B_n$ for every n, and therefore $d(x, y) \leq \operatorname{diam} B_n = \frac{\operatorname{diam} B}{2^n}$. Letting n go to infinity we get $d(x, y) = 0$ and so $x = y$.

Now $x \in B$ implies there exists an $\alpha_x \in A$ so that $x \in G_{\alpha_x}$. But then this implies that there is an $r > 0$ so that $N_r(x) \subseteq G_{\alpha_x}$.

For n large enough we know $B_n \subseteq N_r(x)$. In fact if $\delta_n < r$ then $B_n \subseteq N_r(x)$. Thus since $\delta_n \to 0$ we know $\delta_n < r$ eventually. But then obviously B_n is covered by a finite collection of the $\{G_\alpha\}_{\alpha \in A}$. Oops! The box B must then be compact.

Theorem (Heine-Borel)**.** *A subset K of $\mathbb{R}^d$ is compact if and only if it is closed and bounded.*

Proof. Let's go!

(⇒) We already showed this direction in general metric spaces.

(⇐) If K is bounded then K is contained in some large closed box B which is compact. Therefore K is a closed subset of a compact set. This implies that K is compact (we showed this last time in Hausdorff spaces).

4 Compactness in Metric Spaces

It turns out that being closed and bounded is not sufficient to guarantee compactness in infinite-dimensional metric spaces.

Example. Let $\ell^\infty(\mathbb{N})$ denote the set of bounded sequences $(a_n)_{n\in\mathbb{N}}$. The metric on $\ell^\infty(\mathbb{N})$ is defined as:

$$d((a_n),(b_n)) = \sup_{n\in\mathbb{N}} |a_n - b_n|$$

Consider the set $B = \{(a_n) \in \ell^\infty(\mathbb{N}) \mid \sup_{n\in\mathbb{N}} |a_n| \leq 1\}$.

Exercise. *This set is closed and bounded (check ✓).*

Proof. To note that it's bounded consider that:

$$d((a_n), 0) = \sup_{n\in\mathbb{N}} |a_n| \leq 1$$

So this is trivial. Now consider a sequence of sequences $\left(a_n^{(j)}\right)_{j\in\mathbb{N}}$ which are all in B which converges to some $(a_n)_{n\in\mathbb{N}}$. We will show 1 is an upper bound for the set $\{|a_n|\}_{n\in\mathbb{N}}$, and so:

$$\sup_{n\in\mathbb{N}} |a_n| \leq 1$$

Fix $n \in \mathbb{N}$. Now fix $\varepsilon > 0$. We know there is some large $j \in \mathbb{N}$ so that:

$$d\left(\left(a_n^{(j)}\right), (a_n)\right) = \sup_{n \in \mathbb{N}} \left|a_n - a_n^{(j)}\right| < \varepsilon$$

Now note that:

$$\begin{aligned} |a_n| &\overset{\triangle}{\leq} \left|a_n^{(j)}\right| + \left|a_n - a_n^{(j)}\right| \\ &< 1 + \varepsilon \end{aligned}$$

And so since this holds for all $\varepsilon > 0$ we must have $|a_n| \leq 1$ as desired.

Claim. *This set B is not compact!*

Proof. Consider the sequence of sequences:

$$\left(a_n^{(k)}\right) = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{otherwise} \end{cases}$$

Therefore:

$$d\left(\left(a_n^{(k)}\right), \left(a_n^{(k')}\right)\right) = 1$$

Thus this sequence of sequences $\left(a_n^{(k)}\right)$ can have no convergent subsequence. And thus B is not sequentially compact, and so B is not compact.

How do we fix this? It turns out we need to strengthen our conditions

- Replace closed by Cauchy Complete
- Replace bounded by total boundedness

Definition. *A subset E of a metric space X is* totally bounded *if for every $\varepsilon > 0$ there is a finite cover of E by balls of radius $\varepsilon > 0$.*

Exercise. *Show that:*

- *On $\mathbb{R}^d$ we have boundedness if and only if total boundedness*
  - *Totally bounded implies bounded on every metric space*
  - *For bounded implies totally bounded. Since any box B of the form $[-N,N]^d$ can be split into finitely many subboxes of diameter less than ε, and each sub-box is contained in a ball of radius ε.*
- *On $\mathbb{R}^d$ we have closed if and only if Cauchy complete. Of course Cauchy complete implies closed, and for the other direction we just use Cauchy completeness of $\mathbb{R}^d$.*
- *On $\ell^\infty(\mathbb{N})$ we have that total boundedness is stronger than boundedness. In fact:*

  Exercise. *Show that the set B in the above is bounded but not totally bounded. Use the exact same sequence as in the example and use pigeonhole principle.*

Proof. We've already proved it is bounded. Let $\varepsilon = \frac{1}{2}$ and suppose for the sake of contradiction that we have a finite cover by balls of radius ε. Call these balls $B_1, \dots, B_N$. Without loss of generality assume we have $\left(a_n^{(k)}\right) \in B_j$ for each $1 \leq k \leq N$ where we have:

$$\left(a_n^{(k)}\right) = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{otherwise} \end{cases}$$

Now consider the sequence $\left(a_n^{(m)}\right)$ where we set $m := N+1$. We know there is some k so that $\left(a_n^{(m)}\right) \in B_k$. But then letting $\left(x_n^{(k)}\right)$ be the center of the ball B_k we have that:

$$\begin{aligned} 1 = d\left(\left(a_n^{(m)}\right), \left(a_n^{(k)}\right)\right) &\overset{\triangle}{\leq} d\left(\left(a_n^{(m)}\right), \left(x_n^{(k)}\right)\right) + d\left(\left(x_n^{(k)}\right), \left(a_n^{(k)}\right)\right) \\ &< \frac{1}{2} + \frac{1}{2} = 1 \end{aligned}$$

Oops! We win ☺

Theorem. *Let X be a metric space and $E \subseteq X$. The following are equivalent:*

1) E is compact

2) E is sequentially compact

3) E is complete and totally bounded.

Remark. If X is a complete metric space then 3) above can be replaced by closed and totally bounded.

Lemma. *Completeness of $E \subseteq X$ implies E is closed.*

Proof. Let E be complete and $x_n \in E$ such that $x_n \to x \in X$. Since (x_n) converges it must be Cauchy, and so since E is compelte we know (x_n) converges to some point in E. But limits are unique in metric spaces so $x \in E$, so E is closed!!!

Handout 2

- **Relatively Open, closed, and compact.** We saw in class that the interval $[0, 1)$ is not open in $\mathbb{R}$, but is open relative to the half-line $[0, \infty)$ (taking the usual metric on $\mathbb{R}$ and $[0, \infty)]$). Let us try to formalize and generalize this.

  Let (X, d) be a metric space and $Y \subset X$. Y is a metric space itself, by restricting the metric d to $Y \times Y$.

Q1) Let $E \subset Y$. We say that E is open relative to Y if it is open in the metric space (Y, d). Untangle what this definition means in terms of $N_\delta(p)$ neighborhood of a point $p \in E$ (i.e. restate the condition that E is open in Y in terms of the $N_\delta(p)$ neighborhoods of $p \in E$) and compare it to the condition of E being open in X.

Q2) Deduce that if there is an open subset G of X, then $G \cap Y$ is open relative to Y.

Q3) Show that E is open relative to Y if and only if there exists an open subset G of X such that $E = G \cap Y$.

Q4) Compactness on the other hand behaves better. Suppose that $K \subset Y \subset X$. Show that K is compact relative to X if and only if it is compact relative to Y.

Conclusion: We always need to specify the ambient space when we talk about open/closed sets (that's why we always say "E is an open subset of X"), but we can make statements like "K is compact (or a compact metric space)" without the need to specify the ambient space.

- **The Cantor set.** Let us start with the interval $C = [0, 1]$ and remove the middle third open interval $(\frac{1}{3}, \frac{2}{3})$. This leaves us with the set $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ formed of 2 closed subintervals. Having constructed $C_1 \supset$

$C_2 \supset \ldots \supset C_n$ where C_n is the union of 2^n subintervals each of length $\frac{1}{3^n}$, we construct C_{n+1} as follows: To obtain C_{n+1} we remove the middle third of each of the 2^n intervals that form C_n. This leaves us with a union of 2^{n+1} intervals each of length $\frac{1}{3^{n+1}}$.

Q5) Let $C = \cap_{n=1^\infty} C_n$. Why is C non-empty? Is it compact?

Q6) Show that every point in C is a limit point. Hence C is a perfect set.

Conclusion: From the homework (HW 2), we deduce that C *is uncountable, since any perfect subset of* $\mathbb{R}^d$ *is uncountable.*

Q7) Show that C cannot contain any interval (a, b).

Conclusion: As such, C *is totally disconnected (it has no nontrival connected subset) and nowhere dense (the interior of its closure is empty).*

Q8) What is the total length of C_n? What would be a reasonable definition of the length of C?

MATH 395 Notes

Faye Jackson

September 11, 2020

Proof of Q1. Let $N_\delta^Y(p) = \{a \in Y \mid d(p,a) < \delta\}$ denote the δ-neighborhoods in Y for $p \in Y$ and let $N_\delta(p)$ denote the neighborhood relative to X. Now the definition of an open set $E \subseteq Y$ says that for all $p \in E$ there exists a $\delta > 0$ such that $N_\delta^Y(p) \subseteq E$. Note that:

$$N_\delta^Y(p) = N_\delta(p) \cap Y$$

And so we must have that $N_\delta(p) \cap Y \subseteq E$.

If E were open in X then we would have a stronger condition, namely that the whole neighborhood $N_\delta(p) \subseteq E$.

Proof of Q2. Suppose that G is an open subset of X. Now consider some $p \in G \cap Y$. We know since $p \in G$ that there exists some $\varepsilon > 0$ so that $N_\varepsilon(p) \subseteq G$. But then we know that:

$$N_\varepsilon^Y(p) = N_\varepsilon(p) \cap Y \subseteq G \cap Y$$

By using facts from elementary set theory. This is great! We win now since this must mean that $G \cap Y$ is open as a subset of Y.

Proof of Q3. The backward direction is exactly a consequence of Q2. We work instead on the forward direction.

Suppose that E is open relative to Y. For each $p \in Y$ there exists some $\delta_p > 0$ so that:

$$N_{\delta_p}^Y(p) = N_{\delta_p}(p) \cap Y \subseteq E$$

Now consider the following union:

$$G := \bigcup_{p \in E} N_{\delta_p}(p)$$

Since each $N_{\delta_p}(p)$ is open in X we know that G must be open relative to X. We will show that $E = G \cap Y$.

($\subseteq$) Fix $p \in E$. Then we know that $p \in Y$ since E is a subset of Y, and further we know that $p \in N_{\delta_p}(p)$, and so $p \in G$.

($\supseteq$) Fix $x \in G \cap Y = Y \cap G$. Then:

$$x \in Y \cap G = Y \cap \bigcup_{p \in E} N_{\delta_p}(p) = \bigcup_{p \in E} (Y \cap N_{\delta_p}(p))$$

And thus there exists some p so that:

$$x \in N_{\delta_p}(p) \cap Y = N^Y_{\delta_p}(p) \subseteq E$$

Therefore $x \in E$ just as desired! Great.

With this we win ☺

Proof of Q4. Suppose that $K \subseteq Y \subseteq X$. Now lets go in each direction

($\Rightarrow$) Suppose that K is compact relative to X. Now fix an open cover $\{U_\alpha\}_{\alpha \in A}$ of K relative to Y. By Q3 for each $\alpha \in A$ there exists a G_α which is open in X so that $U_\alpha = G_\alpha \cap Y$. Therefore:

$$\begin{aligned} K &\subseteq \bigcup_{\alpha \in A} U_\alpha = \bigcup_{\alpha \in A} (Y \cap G_\alpha) = Y \cap \bigcup_{\alpha \in A} G_\alpha \\ K &\subseteq \bigcup_{\alpha \in A} G_\alpha \end{aligned}$$

Great! Thus the $\{G_\alpha\}_{\alpha \in A}$ cover K. Since K is compact in X we know there exists a finite subcover $G_{\alpha_1}, \ldots, G_{\alpha_n}$. Then since $K \subseteq Y$ and $K \subseteq \bigcup_{i=1}^n G_{\alpha_i}$ we know:

$$K \subseteq Y \cap \bigcup_{i=1}^n G_{\alpha_i} = \bigcup_{i=1}^n (Y \cap G_{\alpha_i} = \bigcup_{i=1}^n U_{\alpha_i}$$

And therefore $U_{\alpha_1}, \ldots, U_{\alpha_n}$ is a finite subcover of $\{U_\alpha\}_{\alpha \in A}$ just as desired! Great!!

($\Leftarrow$) Suppose that K is compact relative to Y. Now fix an open cover $\{G_\alpha\}_{\alpha \in A}$ of K relative to X. By Q2 we must have that $U_\alpha := G_\alpha \cap Y$ is open in Y for each $\alpha \in A$. Note then that since $K \subseteq Y$ and $K \subseteq \bigcup_{\alpha \in A} G_\alpha$ we know:

$$K \subseteq Y \cap \bigcup_{\alpha \in A} G_\alpha = \bigcup_{\alpha \in A} (Y \cap G_\alpha) = \bigcup_{\alpha \in A} U_\alpha$$

And so $\{U_\alpha\}_{\alpha \in A}$ is an open cover of K in Y. Therefore there must exist a finite subcover for it by compactness, which we will denote by $U_{\alpha_1}, \ldots, U_{\alpha_n}$. Therefore:

$$\begin{aligned} K &\subseteq \bigcup_{i=1}^{n} U_{\alpha_i} = \bigcup_{i=1}^{n} (Y \cap G_{\alpha_i}) = Y \cap \bigcup_{i=1}^{n} G_{\alpha_i} \\ K &\subseteq \bigcup_{i=1}^{n} G_{\alpha_i} \end{aligned}$$

And so $G_{\alpha_1}, \ldots, G_{\alpha_n}$ is a finite subcover of $\{G_\alpha\}_{\alpha \in A}$ just as desired!!!

With this we win ☺

Proof of Q5. For notational convenience denote for $n \in \mathbb{N}_0$:

$$C_n = \bigcup_{i=1}^{2^n} [a_i^n, b_i^n]$$

So that inductively for $1 \leq i \leq 2^n$:

$$\begin{aligned} C_0 &= [0, 1] \\ \left[a_{2i-1}^{n+1}, b_{2i-1}^{n+1}\right] &= \left[a_i^n, \frac{2a_i^n + b_i^n}{3}\right] \\ \left[a_{2i}^{n+1}, b_{2i}^{n+1}\right] &= \left[\frac{a_i^n + 2b_i^n}{3}, b_i^n\right] \end{aligned}$$

Now lets tackle both of these questions!

- Note that $a_1^0 = 0$ will always lie at the edge of an interval because supposeing

$a_1^n = 0$ we know $a_1^{n+1} = a_{2\cdot 1-1}^{n+1} = a_1^n = 0$. Therefore since:

$$0 \in [a_1^n, b_1^n] \subseteq C_n$$

for each $n \geq 0$ we must know that $0 \in C$. A similar argument shows that $1 \in C$.

- C is compact!!! Why? Note that for every $n \geq 0$ we have that C_n is a finite union of closed intervals, so each C_n is closed. Thus, $C = \bigcap_{n=0}^{\infty} C_n$ is closed. Furthermore since $C_0 = [0, 1]$ is closed and bounded, that is compact. Therefore since $C \subseteq C_0$ is a closed subset of a compact set, C must be compact.

Perfect! We win!

Proof of Q6. Fix some point $x \in C$. Then $x \in C_n$ for all $n \geq 0$, and so for each $n \geq 0$ there exists some $1 \leq i_n \leq 2^n$ so that $x \in \left[a_{i_n}^n, b_{i_n}^n\right]$. We claim that $x_n^\ell := a_{i_n}^n$ is a sequence lying in $C \setminus \{x\}$ that converges to x or $x_n^r := a_{i_n}^n$ is a sequence lying in $C \setminus \{x\}$ that converges to x. We tackle this in steps.

- First we show that for all $n \geq 0$ and all $1 \leq i \leq 2^n$ we have a_i^n is in C. First note that $a_i^n \in [a_i^n, b_i^n] \subseteq C_n$, and thus for each $0 \leq m < n$ we must have $a_i^n \in C_n \subseteq C_m$. Inductively we will show that for $m \geq n$ if we let $j_n = i$ and $j_{m+1} = 2j_m - 1$ then:

$$a_i^n = a_{j_m}^m \in C_m$$

  Note that it's trivial for $m = n$. Now suppose that $a_{j_m}^m = a_i^n$. Consider that:

$$a_{j_{m+1}}^{m+1} = a_{2j_m-1}^{m+1} = a_{j_m}^m = a_i^n$$

  And so we must have that this works! Great.

- Now we show that for all $n \geq 0$ and all $1 \leq i \leq 2^n$ we have b_i^n is in C. First note that $b_i^n \in [a_i^n, b_i^n] \subseteq C_n$, and thus for each $0 \leq m < n$ we must have $b_i^n \in C_n \subseteq C_m$. Inductively we will show that for $m \geq n$ if we let $j_n = i$ and $j_{m+1} = 2j_m$ then:

$$b_i^n = b_{j_m}^m \in C_m$$

Note that it's trivial for $m = n$. Now suppose that $b^m_{j_m} = b^n_i$. Consider that:

$$b^{m+1}_{j_{m+1}} = b^{m+1}_{2j_m} = b^m_{j_m} = b^n_i$$

And so we must have that this works! Great.

- Now we show that for each $n \geq 0$ and each $1 \leq i \leq 2^n$ the interval $[a^n_i, b^n_i]$ has length $\frac{1}{3^n}$. TODO

Proof of Q7. TODO

Proof of Q8. TODO

MATH 395 Notes

Faye Jackson

September 14, 2020

Theorem. *Let E be a subset of a metric space X. Then the following are equivalent:*

1) E is compact

2) E is sequentially compact

3) E is complete and bounded.

We've already seen that in matric spaces compactness implies sequential compactness. It remains to show:

(a) Sequential compactness implies compactness

(b) Sequential compactness implies totally bounded and complete

(c) Totally bounded and complete implies sequentially compact

We will prove (b) and (c) first and then (a). In fact, the proof of the theorem follows from the following three lemmas

Lemma 1. *A sequentially compact subset E of X is totally bounded and complete*

Lemma 2. *A totally bounded and complete subset E of X is sequentially compact*

Lemma 3. *A sequentially compact subset of a metric space is compact*

Proof of Lemma 1, Totally Bounded. Note that if $E = \emptyset$ then we are done. Thus let $E \neq \emptyset$ for the duration of this proof.

Let E be sequentially compact. To show it is totally bounded, fix an $\varepsilon > 0$.

Claim. *Let $A \subseteq E$ be a set of points of mutual distance $\geq \varepsilon$. Then A has to be finite*

Proof of claim. Suppose that A were infinite. Then we get a sequence of points $(x_n) \in A$ such that $d(x_n, x_m) \geq \varepsilon$ for all $n \neq m$. But this means that no subsequence of (x_n) is Cauchy, and therefore no subsequence of (x_n) is convergent, violating sequential compactness.

Now let $p_1 \in E$ be arbitrary. If possible we pick $p_2 \in E$ such that $d(p_2, p_1) \geq \varepsilon$. If this is not possible then we stop. Then we pick $p_3 \in E$ such that $d(p_1, p_3) \geq \varepsilon$ and $d(p_2, p_3) \geq \varepsilon$. If this is not possible we stop

Now having picked $p_1, \ldots, p_n$ in this way such that $d(p_i, p_j) \geq \varepsilon$ for all $1 \leq i \neq j \leq n$, we pick $p_{n+1} \in E$ such that $d(p_{n+1}, p_i) \geq \varepsilon$ for all $1 \leq j \leq n$. If this is not possible, then $E \subseteq \bigcup_{i=1}^{n} N_\varepsilon(p_j)$ and we are done.

The claim above tells us that we cannot continue this process forever, and thus it must end after n steps for some $n \in \mathbb{N}$. Therefore E is totally bounded

Proof of Lemma 1, Completeness. Let (x_n) be a Cauchy sequence in E. Since E is sequentially compact there is a convergent sequence (x_{n_k}) such that x_{n_k} converges to some $p \in E$ as k goes to infinity. Now let $\varepsilon > 0$, then there is some $N \in \mathbb{N}$ large enough so that for $k > N$ and $n > N$ we know that:

$$\begin{aligned} d(x_n, x_{n_k}) &< \frac{\varepsilon}{2} \\ d(x_{n_k}, p) &< \frac{\varepsilon}{2} \\ d(x_n, p) &\overset{\triangle}{\leq} d(x_n, x_{n_k}) + d(x_{n_k}, p) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Thus (x_n) converges to $p \in E$. Therefore E is complete!

Proof of Lemma 2. Assume that E is totally bounded and complete. Let (x_n) be a sequence in E. We want to show that it has a convergent subsequence. If the set of all $\{x_n\}$ is finite, then we can find a constant subsequence and we are done. Assume that $\{x_n\}$ is infinite.

Since E is totally bounded, one can cover E with finitely many $\frac{1}{2}$-neighborhoods. One of these neighborhoods must contain infinitely many (x_n) by the pigeonhole principle. Thus we may call this resulting subsequence $\left(x_n^{(1)}\right)$

Now cover E with finitely many $\frac{1}{2^2}$-neighborhoods. One of these neighborhoods contains infinitely many of the $\left(x_n^{(1)}\right)$ by the pigeonhole principle. This gives a subsequence $\left(x_n^{(2)}\right)$ of $\left(x_n^{(1)}\right)$ completely contained in a $\frac{1}{2^2}$-neighborhood. This is also a subsequence of (x_n) of course.

Inductively, we can define a successivesubsequence $\left(x_n^{(k)}\right)$ such that $\left(x_n^{(k)}\right)$ is a subsequence of $\left(x_n^{(k-1)}\right)$ and $\left(x_n^{(k)}\right)$ is contained in a ball of radius $\frac{1}{2^k}$.

Now set $a_n = x_n^{(n)}$. This is a subsequence of (x_n) that satisfies:

$$d(a_n, a_m) = d\left(x_n^{(n)}, x_m^{(m)}\right)$$

If $m \geq n$ then $\left(x_p^{(m)}\right)$ is a subeequence of $\left(x_p^{(n)}\right)$ and $\left(x_p^{(n)}\right)$ is contained in a ball of radius $\frac{1}{2^n}$ with some center, say c for concreteness. Thus:

$$\begin{aligned} d\left(x_n^{(n)}, x_m^{(m)}\right) &\overset{\triangle}{\leq} d\left(x_n^{(n)}, c\right) + d\left(x_m^{(m)}, c\right) \\ &< \frac{1}{2^n} + \frac{1}{2^n} = \frac{1}{2^{n-1}} \end{aligned}$$

Of course we can swap the role of n and m and so we always have:

$$d(a_n, a_m) \leq \frac{1}{2^{\min(n,m)-1}}$$

With this established it is clear that (a_n) is cauchy. By completeness of E, we know (a_n) converges to a point $p \in E$ as desired. Therefore (x_n) has a convergent subsequence

Lemma 4 (3'). *Let $E \subseteq X$ be sequentially compact. Let $\{G_\alpha\}_{\alpha \in A}$ be an open cover of E. Then there exists an $\varepsilon > 0$ such that every ball of radius ε and center $p \in E$ is contained in one of G_α for some $\alpha \in A$.*

Proof. Suppose the statement is not true. Then for any integer $n \geq 1$ there exists a $p_n \in E$ such that $N_{\frac{1}{n}}(p_n)$ is not contained in any of the $\{G_\alpha\}_{\alpha \in A}$. By sequential compactness, (p_n) has a convergent subsequence (p_{n_k}) converging to some $p \in E$.

Since $p \in E$ there exists a α_0 such that $p \in G_{\alpha_0}$, and so there is some $\delta > 0$ so that $N_\delta(p) \subseteq G_{\alpha_0}$.

Since $p_{n_k} \to p$, we may pick n_k large enough so that:

$$d\left(p_{n_k}, p\right) < \frac{\delta}{2} \qquad\qquad \frac{1}{n_k} < \frac{\delta}{2}$$

But then fixing $x \in N_{\frac{1}{n_k}}(p_{n_k})$ we have:

$$d(x,p) \overset{\triangle}{\leq} d\left(x, p_{n_k}\right) + d\left(p_{n_k}, p\right) \qquad\qquad < \frac{\delta}{2} + \frac{\delta}{2} < \delta$$

And so $x \in N_\delta(p) \subseteq G_{\alpha_0}$. This shows that $N_{\frac{1}{n_k}}(p_{n_k}) \subseteq G_{\alpha_0}$. Oops! ☺

Proof of Lemma 3. Suppose that E is sequentially compact. Now let $\{G_\alpha\}$ be any open cover of E. By Lemma 4 (3'), there exists an $\varepsilon > 0$ such that any ε-neighborhood of a point in E is contained in one of the G_α. Since sequentially compact implies totally bounded, E can be covered by finitely many ε-neighborhoods.

That is there is a list $p_1, \ldots, p_N \in E$ such that:

$$E \subseteq \bigcup_{j=1}^{N} N_\varepsilon(p_j)$$

Now for each p_j with $1 \leq j \leq N$ there exists some α_j such that $N_\varepsilon(p_j) \subseteq G_{\alpha_j}$ by construction of ε by lemma 3'. Therefore:

$$E \subseteq \bigcup_{j=1}^{N} G_{\alpha_j}$$

Thus, E is comapct as desired.

MATH 395 Notes

Faye Jackson

September 16, 2020

A Small Digression

Last time we showed that compactness in a metric space is equivalent to sequential compactness is equivalent to totally bounded and complete.

It is clear then that if the total space is complete then compactnes in that space is equivalent to closed and totally bounded

How is this useful in mathematics?

When solving an ODE or a PDE, we can often recast the problem as solving an equation of the form:

$$F(x) = 0$$

for some continuous function $F : X \to X$ and some metric space X, which will be a space of functions. Suppose we are able to find a sequence of approximate solutions to this equation, for example a sequence x_n such that:

$$F(x_n) = \varepsilon_n$$

Where we have $\|\varepsilon_n\|_X \to 0$ as $n \to \infty$. If we can then show that the sequence (x_n) belongs to a compact subset of X, then it must have a convergent subsequence. This convergent subsequence will converge to some x_0, and necessarily we will have $F(x_0) = 0$ as desired.

3 Continuous functions on metric spaces

Definition. *Let X and Y be metric spaces. We say that a function $f : X \to Y$ is continuous at $x_0 \in X$ provided that for every $\varepsilon > 0$ there exists a $\delta = \delta_{\varepsilon,x_0}$ such that whenever $d(y, x_0) < \delta$ we have $d(f(y), f(x_0)) < \varepsilon$.*

In other words, f maps $B_X(x_0, \delta)$ into $B_Y(f(x_0), \varepsilon)$. We say in particular that f is continuous when f is continuous at every point $x_0 \in X$

Proposition. *$f : X \to Y$ is continuous if and only if the inverse image of every open set $U \subseteq Y$ is open in X.*

Proof. Let's go!

($\Rightarrow$) Fix $x \in f^{-1}(U)$. Then since $f(x) \in U$, we know that there is an $\varepsilon > 0$ so that $B_Y(f(x), \varepsilon) \subseteq U$. By continuity there exists some $\delta > 0$ so that f maps $B_X(x, \delta)$ into $B_Y(f(x), \varepsilon$. Therefore:

$$B_X(x, \delta) \subseteq f^{-1}(B_Y(f(x), \varepsilon)) \subseteq f^{-1}(U)$$

Therefore $f^{-1}(U)$ is open.

($\Leftarrow$) Fix $x \in X$. Now fix $\varepsilon > 0$. Note that $B_Y(f(x), \varepsilon)$ is an open set in Y. Thus $f^{-1}(B_Y(f(x), \varepsilon))$ is open in X. Since x is in this set in particular, we know there exists a $\delta > 0$ so that:

$$B_X(x, \delta) \subseteq f^{-1}(B_Y(f(x), \varepsilon))$$
$$f(B_X(x, \delta)) \subseteq B_Y(f(x), \varepsilon)$$

Therefore f is continuous at x. Since $x \in X$ was arbitrary, f is continuous.

Theorem. *Let X be a compact metric space and let $f : X \to Y$ be continuous, then $f(X)$ is compact*

Proof. Let $\{G_\alpha\}$ be an open cover of $f(X)$. Then $\{f^{-1}(G_\alpha)\}$ is an open cover of X. By compactness of X, there exists $\alpha_1, \ldots, \alpha_n$ such that $\{f^{-1}(G_{\alpha_i})\}_{1 \le i \le n}$ is an open cover of X. But then $\{G_{\alpha_i}\}_{1 \le i \le n}$ is an open cover of $f(X)$.

Corrolary 1 (Extreme Value Theorem)**.** *Let $f : X \to \mathbb{R}$ be a continuous function. If f is compact, then f has a maximum and a minumum value.*

Proof. $f(X)$ is compact in $\mathbb{R}$. Therefore $f(X)$ is closed and bounded. Since it is bounded, $\inf f$ and $\sup f$ exist. Furthermore, since it is closed, we know that $\inf f, \sup f \in f(X)$. This shows that these are in fact a minimum and a maximum, as desired.

Definition. *Let X and Y be metric spaces. We say that $f : X \to Y$ is uniformly continuous if for every $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon)$ such that if $d_X(x, y) < \delta$ then $d_Y(f(x), f(y)) < \varepsilon$*

Clearly uniform continuity implies continuity.

Theorem. *Let X be a compact metric space and Y be any metric space. If $f : X \to Y$ is continuous then it is in fact uniformly continuous.*

Proof. Pick some $\varepsilon > 0$. Let $\varepsilon' := \frac{\varepsilon}{2}$. Then for each $x \in X$ we know there is some $\delta_x > 0$ so that $f(B_{\delta_x}(x)) \subseteq B_{\varepsilon'}(f(x))$ by continuity. Let $\delta'_x := \frac{1}{2}\delta_x$. Now note that X is covered by these balls $\{B_{\delta'_x}(x)\}_{x \in X}$. So in particular since X is compact we have $x_1, \dots, x_n$ and $\delta'_1, \dots, \delta'_n > 0$ such that X is covered by $\{B_{\delta'_i}(x_i)\}_{1 \le i \le n}$. Note that we've notated $\delta'_i := \delta'_{x_i}$ and $\delta_i := \delta_{x_i}$ for convenience. Set:

$$\delta := \min_{1 \le i \le n} \delta'_i$$

Now let $x, y \in X$ so that $d(x, y) < \delta$. We know that there is some $1 \le i \le n$ so that $x \in B_{\delta'_i}(x_i)$. Then in particular:

$$\begin{aligned} d(x_i, y) &\overset{\triangle}{\le} d(x_i, x) + d(x, y) < \delta'_i + \delta \\ &\le \delta'_i + \delta'_i = \delta_i \end{aligned}$$

Therefore since $\delta'_i < \delta_i$ it is clear that $x, y \in B_{\delta_i}(x_i)$. Great! Then we must have that $f(x), f(y) \in B_{\varepsilon'}(f(x))$. Which gives:

$$d(f(x), f(y)) \overset{\triangle}{\le} d(f(x), f(x_i)) + d(f(x_i), f(y)) < \varepsilon' + \varepsilon' = \varepsilon$$

Awesome! We win! f is uniformly continuous. See Hani's notes for an equivalent way to do this with Lemma 3' from previous lecture (it is a similar idea).

Part II

Differentiation on $\mathbb{R}^d$

1 Definition of the derviative

1.1 Recollection

Recall. For $\phi : I \to \mathbb{R}$ where I is an open subset of $\mathbb{R}$, we call ϕ differentiable at $x_0 \in I$ provided that the limit

$$\lim_{h\to 0} \frac{\phi(x_0+h)-\phi(x_0)}{h}$$

exists. If so we call this limit $\phi'(x_0)$.

We call ϕ differentiable in I if it is differentiable at every point $x \in I$¿ If I is not open, then we say ϕ is differentiable on I if there exists an extension Φ of ϕ to some open set $J \supseteq I$ such that $\Phi = \phi$ on I and Φ is differentiable on J.

1.2 Generalization Steps

How do we generalize this? We would like to look at functions $\phi : \mathbb{R}^n \to \mathbb{R}^m$ for $n, m \in \mathbb{N}$. If $n = 1$ and $m \geq 1$ then the same definition works:

$$\phi'(x_0) = \lim_{h\to 0} \frac{\phi(x_0+h)-\phi(x_0)}{h}$$

Exercise. *Show that $\phi = (\phi_1, \ldots, \phi_m) : I \to \mathbb{R}^m$ where $I \subseteq \mathbb{R}$ where $I \subseteq \mathbb{R}$ is differentiable at x_0 if and only if ϕ_j is differentiable at x_0 ofor every $1 \leq j \leq m$ and moreover:*

$$\phi'(x_0) = (\phi_1'(x_0), \ldots, \phi_m'(x_0))$$

We run into trouble when $n \geq 1$ we run into trouble because we cannot divide by a vector. Let's reinterpret the case where $n = 1$ to deal with this. Note that:

$$\lim_{h\to 0} \frac{\phi(x_0+h)-\phi(x_0)}{h} - \phi'(x_0) = 0$$
$$\lim_{h\to 0} \frac{\phi(x_0+h)-\phi(x_0)-\phi'(x_0)\cdot h}{h} = 0$$
$$\lim_{h\to 0} \frac{|\phi(x_0+h)-\phi(x_0)-\phi'(x_0)h|}{|h|} = 0$$

The final definition of differntiability at x_0 makes much better since for $\phi : \mathbb{R}^n \to \mathbb{R}^m$, since $|h|$ is a nonzero real number. But we need to properly interpret $\phi'(x_0)h$.

Note that for $\phi : \mathbb{R} \to \mathbb{R}^m$, then $\phi'(x_0)$ provides the best linear approximation to $\phi(x_0+h)-\phi(x_0)$. Namely if $\Delta_h\phi(x_0) = \phi(x_0+h)-\phi(x_0)$ then the definition of $\phi'(x_0)$ tells us that:

$$r(h) := \Delta_h\phi(x_0) - \phi'(x_0)h$$

Satisfies $\frac{|r(h)|}{|h|} \to 0$ as $h \to 0$. Essentially, this means that $\phi'(x_0)h$ takes the increment h in x and gives us the best linear approximation to $\Delta_h\phi(x_0)$. This means that $\phi'(x_0)$ can be interepted as a linear transformation from $\mathbb{R}$ to $\mathbb{R}^m$

1.3 The Correct Generalization

Definition. *Let $E \subseteq \mathbb{R}^n$ be open and let $f : E \to \mathbb{R}^m$. We say that f is differentiable at $x \in E$ provided that there exists a linear transformation $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$ such that:*

$$\lim_{\|h\|\to 0} \frac{\|f(x+h)-f(x)-[Df(x)](h)\|}{\|h\|} = 0$$

We can think of $Df(x_0)$ as an $m \times n$ matrix by linear alegbra. We will prove that $Df(x)$ is unique next lecture, justifying the notation.

Note that the f increment is $\Delta_h f(x) = f(x+h) - f(x)$. How good is the approximation, namely $r(h) = \Delta_h f(x) - Df(x)h$ for a fixed $x \in E$. Then:

$$\lim_{\|h\|\to 0} \frac{\|r(h)\|}{\|h\|} = 0$$

Handout 3

- **The Cantor set.** Let us start with the interval $C = [0,1]$ and remove the middle third open interval $(\frac{1}{3}, \frac{2}{3})$. This leaves us with the set $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ formed of 2 closed subintervals. Having constructed $C_1 \supset C_2 \supset \ldots \supset C_n$ where C_n is the union of 2^n subintervals each of length $\frac{1}{3^n}$, we construct C_{n+1} as follows: To obtain C_{n+1} we remove the middle third of each of the 2^n intervals that form C_n. This leaves us with a union of 2^{n+1} intervals each of length $\frac{1}{3^{n+1}}$.

Q1) Let $C = \cap_{n=1\infty} C_n$. Why is C non-empty? Is it compact?

Q2) Show that every point in C is a limit point. Hence C is a perfect set.

Conclusion: From the homework (HW 2), we deduce that C is uncountable, since any perfect subset of $\mathbb{R}^d$ is uncountable.

Q3) Show that C cannot contain any interval (a, b).

Conclusion: As such, C is totally disconnected (it has no non-trival connected subset) and nowhere dense (the interior of its closure is empty).

Q4) What is the total length of C_n? What would be a reasonable definition of the length of C?

- **Wish list for a measure function** Motivated by the above, it would be grand to have a measure function that tells us how big or small a subset of $\mathbb{R}^d$ is. This would be a function from the set of subsets of $\mathbb{R}^d$ into $[0, \infty]$, say $m : \mathcal{P}(\mathbb{R}^d) \to [0, \infty]$. We would like this function to satisfy the following properties:

a) If $E_1, E_2, \ldots$ is a countable collection of disjoint subsets of $\mathbb{R}$, then

$$m(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} m(E_n).$$

b) If E is congruent to F (i.e. F can be obtained from E by applying rigid motions: translations, rotations, or a reflections) then we should have that $m(E) = m(F)$.

c) $m([0,1)^d) = 1$.

The bad news is that no such function can exist, and here's why (at least when $d = 1$). Let us define an equivalence relation between elements of $[0,1)$ as follows: We say $x \sim y$ if $x - y$ is a rational number. Let N be the subset of $[0,1]$ that contains exactly one element of each equivalence relation (the existence of this N requires invoking the axiom of choice). Now let $R = [0,1) \cap \mathbb{Q}$, and for each $r \in R$ define the set

$$N_r = \{x + r : x \in N \cap [0, 1-r]\} \cup \{x + r - 1 : x \in N \cap [1-r, 1)\}.$$

(Basically N_r is just the translate of N by r units to the right, except that we move the part that sticks out of the interval $[0,1)$ one unit to the left).

Q5) Show that $[0,1)$ is the disjoint union of N_r for $r \in R$.

Q6) Show that if a measure function satisfying a), b) and c) above exists, then $m(N) = m(N_r)$ for every $r \in R$.

Q7) Arrive at a contradiction.

Remark: One might think that possibly relaxing condition a) to cover only *finitely* many disjoint sets E_n, i.e.

$$m(\cup_{n=1}^{N} E_n) = \sum_{n=1}^{N} m(E_n).$$

would resolve the contradiction. Unfortunately, the Banach-Tarski paradox (cf. Figure 1) tells us that this is not enough to resolve this issue.

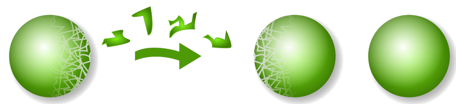


Figure 1: Banach-Tarski tells us that we can split the unit ball in $\mathbb{R}^3$ into finitely many (actually 5 is sufficient) many disjoint pieces, apply rigid motions to those pieces and then reassemble them to obtain two copies of the unit ball.

Conclusion: The problem with the above wishlist is that we insisted on being able to measure *every* subset of $\mathbb{R}^d$. We have shown that this is impossible. The solution is to be content with a measure function that is defined on some but not all subsets. Such subsets will be called measurable subsets.

MATH 395 Notes

Faye Jackson

September 18, 2020

Proof of Q1. For notational convenience denote for $n \in \mathbb{N}_0$:

$$C_n = \bigcup_{i=1}^{2^n} [a_i^n, b_i^n]$$

So that inductively for $1 \leq i \leq 2^n$:

$$C_0 = [0,1]$$
$$\left[a_{2i-1}^{n+1}, b_{2i-1}^{n+1}\right] = \left[a_i^n, \frac{2a_i^n + b_i^n}{3}\right]$$
$$\left[a_{2i}^{n+1}, b_{2i}^{n+1}\right] = \left[\frac{a_i^n + 2b_i^n}{3}, b_i^n\right]$$

Now lets tackle both of these questions!

- Note that $a_1^0 = 0$ will always lie at the edge of an interval because supposeing $a_1^n = 0$ we know $a_1^{n+1} = a_{2\cdot 1-1}^{n+1} = a_1^n = 0$. Therefore since:

  $$0 \in [a_1^n, b_1^n] \subseteq C_n$$

  for each $n \geq 0$ we must know that $0 \in C$. A similar argument shows that $1 \in C$.

- C is compact!!! Why? Note that for every $n \geq 0$ we have that C_n is a finite union of closed intervals, so each C_n is closed. Thus, $C = \bigcap_{n=0}^{\infty} C_n$ is closed. Furthermore since $C_0 = [0,1]$ is closed and bounded, that is compact. Therefore since $C \subseteq C_0$ is a closed subset of a compact set, C must be compact.

Perfect! We win!

Proof of Q2. Fix some point $x \in C$. Then $x \in C_n$ for all $n \geq 0$, and so for each $n \geq 0$ there exists some $1 \leq i_n \leq 2^n$ so that $x \in \left[a^n_{i_n}, b^n_{i_n}\right]$. Suppose that $\varepsilon > 0$, then there is some $N \in \mathbb{N}$ so that $\frac{1}{3^N} < \varepsilon$. We claim that $a^N_{i_N}, b^N_{i_N} \in N_\varepsilon(x) \cap C$

- First we show that for all $n \geq 0$ and all $1 \leq i \leq 2^n$ we have a^n_i is in C. First note that $a^n_i \in [a^n_i, b^n_i] \subseteq C_n$, and thus for each $0 \leq m < n$ we must have $a^n_i \in C_n \subseteq C_m$. Inductively we will show that for $m \geq n$ if we let $j_n = i$ and $j_{m+1} = 2j_m - 1$ then:

  $$a^n_i = a^m_{j_m} \in C_m$$

  Note that it's trivial for $m = n$. Now suppose that $a^m_{j_m} = a^n_i$. Consider that:

  $$a^{m+1}_{j_{m+1}} = a^{m+1}_{2j_m - 1} = a^m_{j_m} = a^n_i$$

  And so we must have that this works! Great.

- Now we show that for all $n \geq 0$ and all $1 \leq i \leq 2^n$ we have b^n_i is in C. First note that $b^n_i \in [a^n_i, b^n_i] \subseteq C_n$, and thus for each $0 \leq m < n$ we must have $b^n_i \in C_n \subseteq C_m$. Inductively we will show that for $m \geq n$ if we let $j_n = i$ and $j_{m+1} = 2j_m$ then:

  $$b^n_i = b^m_{j_m} \in C_m$$

  Note that it's trivial for $m = n$. Now suppose that $b^m_{j_m} = b^n_i$. Consider that:

  $$b^{m+1}_{j_{m+1}} = b^{m+1}_{2j_m} = b^m_{j_m} = b^n_i$$

  And so we must have that this works! Great.

- Now we show that for each $n \geq 0$ and each $1 \leq i \leq 2^n$ the interval $[a^n_i, b^n_i]$ has length $\frac{1}{3^n}$. Note first that:

  $$b^0_1 - a^0_1 = 1 - 0 = 1 = \frac{1}{3^0}$$

Inductively for $1 \leq i \leq 2^n$ then we know that:

$$b_{2n-i}^{n+1} - a_{2n-i}^{n+1} = \frac{2a_i^n + b_i^n}{3} - a_i^n = \frac{b_i^n - a_i^n}{3} = \frac{1}{3} \cdot \frac{1}{3^n} = \frac{1}{3^{n+1}}$$
$$b_{2n}^{n+1} - a_{2n}^{n+1} = b_i^n - \frac{a_i^n + 2b_i^n}{3} = \frac{b_i^n - a_i^n}{3} = \frac{1}{3} \cdot \frac{1}{3^n} = \frac{1}{3^{n+1}}$$

Now we're done, since in particular $a_{i_N}^N$ and $b_{i_N}^N$ are distinct, so for any ε neighborhood of x there are at least two points in $N_\varepsilon(x) \cap C$. Thus x is a limit point.

Proof of Q3. Fix $a < b$. But then if we had two points $x, y \in (a, b)$ such that $x, y \in C$ and $y > x$. Note that we then know that there exists some $N \in \mathbb{N}$ so that $\frac{1}{3^N} < \varepsilon$. This means that x and y must lie in different intervals making up C_N, since these are disjoint. But then $(a, b) \cap C$ is not an interval, since $x, y \in C \cap (a, b)$ but there is some point z between x and y so that $z \notin C$. This necessarily means so then $(a, b) \neq C \cap (a, b)$, and so $(a, b) \nsubseteq C$.

Proof of Q4. Note that the total length of C^n is:

$$\ell(C_n) = \frac{2^n}{3^n}$$

Since C_n is a union of 2^n disjoint intervals each of length 3^n. Note that for each $n \in \mathbb{N}$ we must conclude since $C \subseteq C_n$ we know:

$$\ell(C) \leq \ell(C_n) = \frac{2^n}{3^n} = \left(\frac{2}{3}\right)^n$$

Taking $n \to \infty$ we then can see that $\ell(C)$ should be zero.

Proof of Q5. Fix $r, q \in R = [0, 1) \cap \mathbb{Q}$. We will first show that if $N_r \cap N_q \neq \emptyset$ then $r = q$, so by contrapositive the $\{N_r\}_{r \in R}$ are disjoint. Fix $y \in N_r \cap N_q$. There are four cases:

- $y = x_r + r$ and $y = x_q + q$ for some $x_r, x_q \in N$. Then $x_r - x_q = q - r$ by algebra, and so since $r, q \in \mathbb{Q}$ we have that $q - r \in \mathbb{Q}$ and so $x_r \sim x_q$. By the definition of N it follows that $x_r = x_q$. Therefore $x_r + r = x_r + q$, giving that $r = q$.

- $y = x_r + r - 1$ and $y = x_q + q - 1$ for some $x_r, x_q \in N$. Then $x_r - x_q = q - r$ by algebra, and so since $r, q \in \mathbb{Q}$ we have that $q - r \in \mathbb{Q}$ and so $x_r \sim x_q$. By the definition of N it follows that $x_r = x_q$. Therefore $x_r + r - 1 = x_r + q - 1$, giving that $r = q$.

- $y = x_r + r - 1$ and $y = x_q + q$ for some $x_r, x_q \in N$. Then $x_r - x_q = q - r + 1 \in \mathbb{Q}$. Thus $x_r = x_q$. Therefore $q = r - 1$ by some quick algebra. This is clearly a contradiction! Why? Well $0 \leq r < 1$, and so $-1 \leq r - 1 < 0$, but we know $q \geq 0$!!! Oops!

- $y = x_r + r$ and $y = x_q + q - 1$ for some $x_r, x_q \in N$. Then $x_r - x_q = q - r - 1 \in \mathbb{Q}$. Thus $x_r = x_q$. Therefore $r = q - 1$ by some quick algebra. This is clearly a contradiction! Why? Well $0 \leq q < 1$, and so $-1 \leq q - 1 < 0$, but we know $r \geq 0$!!! Oops!

We want to show that:

$$[0, 1) = \bigcup_{r \in R} N_r$$

Let's go!

($\subseteq$) Fix $y \in [0, 1)$. Then by definition there is some $x \in N$ so that $y \sim x$. Note that then $y - x \in \mathbb{Q}$. Further we have $0 \leq x, y < 1$ There are two cases:

- Suppose that $y - x \geq 0$. Now set $r := y - x$. First note that since $x \geq 0$ and $y < 1$ we know $y - x < 1 - 0 = 1$. Therefore $r \in \mathbb{Q} \cap [0, 1) = R$. We claim that $y \in N_r$. In particular note that $y = x + r$. All that remains to be shown is $x \in [0, 1 - r)$. We know since $x \in N$ that $x \in [0, 1)$, so $x \geq 0$ immediately. We merely need to show that $x < 1 - y + x$. This is simple, since $y < 1$ we know $1 - y > 0$. With this we must have that $x \in [0, 1) \cap N$, and so:

  $$y \in \{x' + r \mid x' \in N \cap [0, 1 - r)\} \subseteq N_r$$

  And so $y \in N_r$

- Suppose that $y - x < 0$. Set $r := y - x + 1$. Note then that $r < 1$. Since $0 \leq y$ we know $-x \leq y - x$, and then since $x < 1$ it follows that $-1 < -x \leq y - x$, and so $0 < r$. This shows since $r \in \mathbb{Q}$ that $r \in R = [0, 1) \cap \mathbb{Q}$. We claim that $y \in N_r$. Note in particular that $y = x + r - 1$ by algebra. We need merely show that $x \in [1 - r, 1)$. To do this note that $y \geq 0$ so $y \leq 0$:

$$x \geq -y + x = 1 - y + x - 1 = 1 - (y - x + 1) = 1 - r$$

And we already know $x < 1$. Therefore:

$$y \in \{x' + r - 1 \mid x' \in N \cap [1 - r, 1)\} \subseteq N_r$$

And so $y \in N_r$!

Great! Since in either case $r \in R$, we must have that $y \in \bigcup_{r \in R} N_r$. This finishes this direction!

($\supseteq$) This side follows fairly immediately. Fix $y \in \bigcup_{r \in R} N_r$. Then $y \in N_r$ for some $r \in R$. There are then two quick cases:

- We have that $y = x + r$ for some $x \in N \cap [0, 1 - r)$. Then note that since $r \geq 0$ we have:

$$\begin{aligned} &0 \leq x < 1 - r \\ &0 \leq r \leq x + r = y < 1 \end{aligned}$$

  And thus $y \in [0, 1)$

- We have that $y = x + r - 1$ for some $x \in N \cap [1 - r, 1)$. Then note that since $1 - r \leq x < 1$ that $-r \leq x - 1 < 0$. Therefore since $r < 1$ we know:

$$0 \leq x + r - 1 < r < 1$$

With this we're done!

We've finished the proof that this is a disjoint union! Wow!

Proof of Q6. Fix some $r \in R$. We wish to show that $m(N_r) = m(N)$. First note that:

$$\begin{aligned} m(N_r) &= m\left(\{x + r \mid x \in N \cap [0, 1 - r)\} \cup \{x + r - 1 \mid x \in N \cap [1 - r, 1)\right) \\ &= m(\{x + r \mid x \in N \cap [0, 1 - r)\}) + m(\{x + r - 1 \mid x \in N \cap [1 - r, 1)\}) \end{aligned}$$

This follows from axiom (a) for our measure. But then by axiom (b) note that these are translations of $N \cap [0, 1-r)$ and $N \cap [1-r, 1)$ respectively so:

$$m(N_r) = m(N \cap [0, 1-r)) + m(N \cap [1-r, 1))$$

We need to now show that:

$$[0, 1) = [0, 1-r) \cup [1-r, 1)$$

This is farily quick since we note that $r \in [0, 1)$

($\subseteq$) Fix $x \in [0, 1)$. Then if $x < 1-r$ we have $x \in [0, 1-r)$. Otherwise we know $x \geq 1-r$ and so $x \in [1-r, 1)$.

($\supseteq$) Fix $x \in [0, 1-r)$. Then since $r \geq 0$ we know $x < 1-r \leq 1$. Therefore $0 \leq x < 1$, and so $x \in [0, 1)$

In the other case, fix $x \in [1-r, 1)$. Then we know since $r < 1$ that $0 < 1-r \leq x$. Therefore since $0 < x < 1$ we have $x \in [0, 1)$.

Now consider that:

$$(N \cap [0, 1-r)) \cup (N \cap [1-r, 1)) = N \cap ([0, 1-r) \cup [1-r, 1)) = N \cap [0, 1) = N$$

The last equality holds sincec N is a subset of $[0, 1)$. Therefore:

$$m(N_r) = m(N \cap [0, 1-r)) + m(N \cap [1-r, 1)) = m(N)$$

And we are done!

Proof of Q7. We wish to arrive at a contradiction. There are three quick cases:

- Suppose that $m(N) = 0$. Then since $\mathbb{Q}$ is countable we know $R = \mathbb{Q} \cap [0, 1)$ is countable, giving us by axiom (a) and (c) that:

$$1 = m([0, 1)) = m\left(\bigcup_{r \in R} N_r\right) = \sum_{r \in R} m(N_r) = \sum_{r \in R} 0 = 0$$

  This is a clear contradiction! Oops!

- Suppose that $m(N) > 0$. Note that R is countable and for $n \geq 2$ we have $0 < \frac{1}{n} < 1$ and so $\frac{1}{n} \in R$. Then using axiom (a), axiom (c), and the fact that $m(N)$ is positive we know that:

$$\begin{aligned} 1 = m([0,1)) &= m\left(\bigcup_{r \in R} N_r\right) \\ &= \sum_{r \in R} m(N_r) \geq \sum_{n=2}^{\infty} m\left(N_{\frac{1}{n}}\right) \\ &= \sum_{n=2}^{\infty} m(N) = \infty \end{aligned}$$

  This is clearly true, since we know that $m(N) > 0$ doesn't go to zero, $\sum_{n=2}^{\infty} m(N)$ must diverge to infinity. This is an oops since $1 < \infty$

- Suppose that $m(N) = \infty$ Then since R is countable and $0 \in R = \mathbb{Q} \cap [0,1)$ we know that by axiom (b) and axiom (c),

$$1 = m([0,1)) = m\left(\bigcup_{r \in R} N_r\right) = \sum_{r \in R} m(N_r) \geq m(N_0) = m(N) = \infty$$

  This cannot be true since $1 < \infty$. Oops!

With all three of these completed, we must conclude that $m(N)$ is undefined!!! Wow! This is amazing ☺

MATH 395 Notes

Faye Jackson

September 21, 2020

Continue Differentiability in higher dimensions

We first recalled the definition of the derivative for $\phi : \mathbb{R} \to \mathbb{R}^d$:

$$\phi'(x) = \lim_{h \to 0} \frac{\phi(x+h) - \phi(x)}{h}$$

But we cannot divide by h if $h \in \mathbb{R}^d$. We reinterpreted the definition saying that $\phi'(x)$ exists if and only if:

$$\lim_{h \to 0} \frac{|\phi(x+h) - \phi(x) - \phi'(x)h|}{|h|} = 0$$

Reinterpreting this for $\phi : \mathbb{R}^n \to \mathbb{R}^m$ we ask for a linear transformation $D\phi(x) : \mathbb{R}^n \to \mathbb{R}^m$ such that:

$$\lim_{h \to 0} \frac{\|\phi(x+h) - \phi(x) - D\phi(x) \cdot h\|}{\|h\|}$$

This recalls the best linear approximation interpretation of the derivative. If we write:

$$\begin{aligned} \Delta\phi(h) &= \phi(x+h) - \phi(x) \\ r(h) &= \Delta\phi(h) - D\phi(x) \cdot h \end{aligned}$$

Then we ask for $\frac{\|r(h)\|}{\|h\|} \to 0$ as $h \to 0$. We write this as $\|r(h)\| = o(\|h\|)$ That is $\|r(h)\| \ll \|h\|$ as $h \to 0$.

Definition. *Let $E \subseteq \mathbb{R}^n$ be open and let $f : E \to \mathbb{R}^m$. We say that f is differentiable at $x \in E$ provided that there is a linear transformation $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$*

such that:

$$\lim_{\|h\|\to 0} \frac{\|f(x+h) - f(x) - Df(x)\cdot h\|}{\|h\|} = 0 \qquad (**)$$

We can of course think of $Df(x)$ as an $m\times n$ matrix. If f is differentiable at every $x \in E$ we say that f is differentiable in E. In this case we have the total derivative:

$$Df : E \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$$

Remark. We have some comments

- We need x to be an interior point of E so that $x+h \in E$ for small h, so that $f(x+h)$ makes sense. When E is open this is automatic.
- The numerator in the difference quotient above is in $\mathbb{R}^m$ whereas the denominator is in $\mathbb{R}^n$.
- Defining $r(h) = f(x+h) - f(x) - Df(x)\cdot h$, we have that $r(h) = o(h)$. That is:

  $$\lim_{h\to 0} \frac{\|r(h)\|}{\|h\|} = 0$$

  Note then that $Df(x)\cdot h = \mathcal{O}(h)$, that is there is a constant $C \in \mathbb{R}$ so that $\|Df(x)\cdot h\| \leq C\|h\|$, but this is different than $r(h) = o(h)$.
- This definition of derivative only makes sense if $Df(x)$ is unique when it exists.

Proposition 1. Let E, f, and $x \in E$ be as in the above definition. Suppose that A_1 and A_2 are two linear transformations such that $(**)$ holds. Then $A_1 = A_2$

Proof. Let $r_j(h) = f(x+h) - f(x) - A_j h$ for $j = 1, 2$. Then we have that $\frac{\|r_j(h)\|}{\|h\|} \to 0$. Let $u \in \mathbb{R}^n$ be arbitrary and nonzero and take $h = tu$ for $t > 0$, then we can divide by $\|tu\|$ to get:

$$\begin{aligned} r_1(tu) - r_2(tu) &= (A_2 - A_1)(tu) = t(A_2 - A_1)u \\ \frac{\|(A_2 - A_1)u}{\|u\|} &= \frac{\|r_1(tu) - r_2(tu)\|}{t\|u\|} \\ &\leq \frac{\|r_1(tu)\|}{\|tu\|} + \frac{\|r_2(tu)\|}{\|tu\|} \end{aligned}$$

Thus $\frac{\|(A_2-A_1)u\|}{\|u\|} \to 0$ as $t \to 0$. Therefore $(A_2 - A_1)u = 0$, so $A_1 u = A_2 u$. Note that clearly $A_1 \cdot 0 = A_2 \cdot 0$. Taking these together we know $A_1 = A_2$.

Example. Let $f(x) = a + Bx$ where $a \in \mathbb{R}^m$ and and $B \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ where $f : \mathbb{R}^n \to \mathbb{R}^m$. Then to compute $Df(x)$ note that:

$$\begin{aligned} f(x+h) - f(x) &= Bh \\ f(x+h) - f(x) - Bh &= 0 \end{aligned}$$

Therefore we know clearly that:

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - Bh\|}{\|h\|} = 0$$

Therefore $Df(x) = B$ for any $x \in \mathbb{R}^n$.

Remark. Of course, if f is differentiable at x, then it must be continuous there. Why? Continuity is equivalent to $\|f(x+h) - f(x)\| \to 0$ as $h \to 0$. Differentiability is equivalent to $\|f(x) + f(x) - Df(x)h\| = \|r(h)\| = o(\|h\|)$. In particular this implies that:

$$\begin{aligned} \|f(x+h) - f(x)\| &= \|Df(x)h + r(h)\| \\ &\overset{\triangle}{\leq} \|Df(x)h\| + \|r(h)\| \end{aligned}$$

But both of these go to 0 as $h \to 0$. Therefore:

$$\lim_{h \to 0} \|f(x+h) - f(x)\| = 0$$

Directional and Partial Derivatives, computing the derivative

Definition. *Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \to \mathbb{R}^m$. Suppose $x \in A$ and $u \in \mathbb{R}^n$ with $u \neq 0$. We define the directional derivative $D_u f(x)$ as the limit:*

$$D_u f(x) := \lim_{t \to 0} \frac{f(x+tu) - f(x)}{t} \in \mathbb{R}^m$$

$$D_u f(x) := \lim_{t \to 0} \frac{f(x+tu) - f(x)}{t} \in \mathbb{R}^m$$

Note that this just means that:

$$D_u f(x) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x+tu)$$

Example. Let $f : \mathbb{R}^2 \to \mathbb{R}$ bed efined as $\sin(x_1 x_2)$. Then let $u = (1, 0)$:

$$\begin{aligned}
D_u f(x_1, x_2) &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \sin((x_1 + t)x_2) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \sin(x_1 x_2 + t x_2) \\
&= (\cos(x_1 x_2 + t x_2) \cdot x_2]_{t=0} \\
&= \cos(x_1 x_2) \cdot x_2
\end{aligned}$$

Theorem. *Let $A \subseteq \mathbb{R}^n$ be open and $f : A \to \mathbb{R}^m$ be differentiable at $x \in A$. THen all directional derivatives $D_u f(x)$ exist at x_0 and:*

$$D_u f(x) = Df(x) \cdot u$$

In particular $D_u f(x)$ is linear in u.

Proof. From the definition of $Df(x)$ we have for any $u \in \mathbb{R}^n \setminus \{0\}$

$$\lim_{t \to 0} \frac{\|f(x+tu) - f(x) - Df(x) \cdot tu\|}{\|tu\|} = 0$$
$$\lim_{t \to 0} \frac{\|f(x+tu) - f(x) - t \cdot (Df(x) \cdot u)\|}{\|tu\|} = 0$$

This implies that:

$$f(x+tu) - f(x) - t \cdot Df(x) \cdot u = r(tu)$$

Therefore $\frac{\|r(tu)\|}{\|tu\|} \to 0$ as $t \to 0$. Dividing by t we get that:

$$\frac{f(x+tu) - f(x)}{t} - Df(x)u = \frac{r(tu)}{t}$$

Therefore:

$$\left\| \frac{f(x+tu) - f(x)}{t} - Df(x)u \right\| = \frac{\|r(tu)\|}{\|t\|} = \|u\| \cdot \frac{\|r(tu)\|}{\|tu\|} \to 0$$

As $t \to 0$. Therefore:

$$\lim_{t\to 0} \left\| \frac{f(x+tu)-f(x)}{t} - Df(x)u \right\| = 0$$
$$D_u f(x) = \lim_{t\to 0} \frac{f(x+tu)-f(x)}{t} = Df(x)\cdot u$$

<u>Caution</u> We will see next time that the converse is not true. Namely, the directional derivatives might exist at x without f being differntiable at x. In that case $D_u f(x)$ might not even be a linear function of u.

Partial Derivatives

Since $D_u f(x) = Df(x)\cdot u$, we can determine $Df(x)$ by letting u range over the standard basis vectors.

Definition. *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ where A is open. The j-th partial derivative of f at x is defined as:*

$$\frac{\partial f}{\partial x_j}(x) = D_{e_j} f(x) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x+te_j)$$

Example. When $m = 1$ we know $f : \mathbb{R}^n \to \mathbb{R}$ then:

$$\begin{aligned}\frac{\partial f}{\partial x_j}(x_1,\dots,x_n) &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x_1,\dots,x_j+t_j,\dots,x_n) \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=x_j} f(x_1,\dots,x_{j-1},s,x_{j+1},\dots,x_n) \\ &= \phi'(x_j)\end{aligned}$$

Where $\phi(s) = f(x_1,\dots,x_{j-1},s,x_{j+1},\dots,x_n)$. This just means that $\frac{\partial f}{\partial x_j}$ is computed by pretending that $x_1,\dots,x_{j-1},x_{j+1},\dots,x_n$ are constant and differentiating with respect to x_j.

MATH 395 Notes

Faye Jackson

September 23, 2020

Recall. We defined directional derivatives for $f : \mathbb{R}^n \to \mathbb{R}^m$ for $u \in \mathbb{R}^n$ by:

$$D_u f(x) = \lim_{t \to 0} \frac{f(x+tu) - f(x)}{t} = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x+tu) \in \mathbb{R}^m$$

We also defined partial derivatives $\frac{\partial f}{\partial x_i} \in \mathbb{R}^m$ for $1 \le i \le n$ by:

$$\frac{\partial f}{\partial x_i} = D_{e_i} f(x)$$

Furthermore, if f is differentiable at x then $D_u f(x)$ exists for every u. Morevoer:

$$D_u f(x) = Df(x) \cdot u$$

The converse is not true in general!!! We will give today an example where $D_u f(x)$ exists for every n but $Df(x)$ does not

If $n \ge 1$ and $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ then we can write $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$. Then for every $u \in \mathbb{R}^n$ we have:

$$\begin{aligned} D_u f(x) = \lim_{t \to 0} \frac{f(x+tu) - f(x)}{t} &= \lim_{t \to 0} \begin{pmatrix} \frac{f_1(x+tu) - f_1(x)}{t} \\ \vdots \\ \frac{f_m(x+tu) - f_m(x)}{t} \end{pmatrix} \\ &= \begin{pmatrix} \lim_{t \to 0} \frac{f_1(x+tu) - f(x)}{t} \\ \vdots \\ \lim_{t \to 0} \frac{f_m(x+tu) - f(x)}{t} \end{pmatrix} = \begin{pmatrix} D_u f_1(x) \\ \vdots \\ D_u f_m(x) \end{pmatrix} \end{aligned}$$

That is, directional derivatives can be taken componentwise. In particular:

$$\frac{\partial f}{\partial x_i} = \begin{pmatrix} \frac{\partial f_1}{\partial x_i} \\ \vdots \\ \frac{\partial f_m}{\partial x_i} \end{pmatrix}$$

Example. Let $F : \mathbb{R}^2 \to \mathbb{R}^3$ be given by:

$$F(x, y) = \begin{pmatrix} x^2 + y^2 \\ xy \\ \sin y \end{pmatrix}$$

Then computing:

$$\begin{aligned} \frac{\partial F}{\partial x}(x, y) &= \begin{pmatrix} \frac{\partial}{\partial x}(x^2 + y^2) \\ \frac{\partial}{\partial x}(xy) \\ \frac{\partial}{\partial x} \sin y \end{pmatrix} = \begin{pmatrix} 2x \\ y \\ 0 \end{pmatrix} \\ \frac{\partial F}{\partial y}(x, y) &= \begin{pmatrix} 2y \\ x \\ \cos y \end{pmatrix} \end{aligned}$$

If $u = (1, 2)$. Then:

$$\begin{aligned} D_u F(x, y) &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \begin{pmatrix} (x + t)^2 + (y + 2t)^2 \\ (x + t)(y + 25) \\ \sin(y + 2t \end{pmatrix} = \begin{pmatrix} 2x + 4y \\ y + 2x \\ 2\cos(y) \end{pmatrix} \\ D_u F(x, y) &= D_{e_1 + 2e_2} F(x, y) = D_{e_1} F(x, y) + 2 D_{e_2} F(x, y) \\ &= \frac{\partial F}{\partial x}(x, y) + 2 \frac{\partial F}{\partial y}(x, y) \end{aligned}$$

This suggests that F is differentiable at (x, y). But it's not a proof.

Theorem. *Let $f : A \subseteq \mathbb{R} \to \mathbb{R}^m$ where A is open and suppose $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$. Then:*

a) f is differentiable $x \in A$ if and only if each of the components of $f_1, \dots, f_n$ are differentiable at x

b) If f is differentiable at $x \in A$, then $Df(x)$ is the $(m \times n)$ matrix whose j-th *column is* $\frac{\partial f}{\partial x_j}$.

c) *Equivalently,* $Df(x)$ *is the* $(m \times n)$ *matrix whose* i*-th row is* $Df_i(x)$.

d) *Equivalently* $Df(x)$ *is the* $m \times n$ *matrix whose* (i,j)*-th entry is* $\frac{\partial f_i}{\partial x_j}(x)$.

Remark. In calculus for $f : \mathbb{R}^n \to \mathbb{R}$, $Df(x) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)$, is often denoted $\nabla f(x)$, the gradient of f at x. Sometimes it is important to distinguish between $Df(x)$ which is a $(1 \times n)$ matrix and $\nabla f(x)$ which is an $(n \times 1)$ matrix, that is a vector.

Proof. f is differentiable at x if and only if there exists an $(m \times n)$ matrix A such that:

$$\frac{\|f(x+h) - f(x) - Ah\|}{\|h\|} \to 0$$

as $\|h\| \to 0$. This holds if and only if each coordinate:

$$\frac{\|f_i(x+h) - f_i(x) - A_i \cdot h\|}{\|h\|} \to 0$$

As $\|h\| \to 0$, where A_i is the i-th row of A. Since the i-th coordinate of Ah is $A_i h$. But this is equivalent to saying that f_i is differentiable at x, and $Df_i(x)$ is equal to the i-th row of $Df(x)$.

The above implies parts a) and c). To obtain part b) and c) note that if f is differentiable at x then:

$$D_u f(x) = Df(x) \cdot u$$

Taking $u = e_j$ for $1 \le j \le n$. we get:

$$\frac{\partial f}{\partial x_j}(x) = Df(x) \cdot e_j$$

But this is exactly the j-th column of $Df(x)$. Therefore:

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Example. Let $F : \mathbb{R}^2 \to \mathbb{R}^3$ to be as before:

$$F(x, y) = \begin{pmatrix} x^2 + y^2 \\ xy \\ \sin y \end{pmatrix}$$

Then if the derivative exists we know:

$$Df(x, y) = \begin{pmatrix} \nabla(x^2 + y) \\ \nabla(xy) \\ \nabla(\sin y) \end{pmatrix} = \begin{pmatrix} 2x & 2y \\ y & x \\ 0 & \cos y \end{pmatrix}$$

Great

Remark. Partial derivatives and even directional derivatives of a function can exist at x even if each f_n is not differentiable at x. Take $f : \mathbb{R}^2 \to \mathbb{R}$:

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^4 + y^2} & \text{if } (x, y) \neq 0 \\ 0 & \text{if } (x, y) = 0 \end{cases}$$

For $u \in \mathbb{R}^2 \setminus \{0\}$ let us compute $D_u f(0)$. Take $u = (u_1, u_2)$. Then:

$$\begin{aligned} D_u f(0) &= \lim_{t \to 0} \frac{f(0 + tu) - f(0)}{t} = \lim_{t \to 0} \frac{t^3 u_1^2 u_2}{t^5 u_1^4 + t^3 u_2^2} \\ &= \lim_{t \to 0} \frac{u_1^2 u_2}{t^2 u_1^4 + u_2^2} = \begin{cases} 0 & \text{if } u_2 = 0 \\ \frac{u_1^2}{u_2} & \text{if } u_2 \neq 0 \end{cases} \end{aligned}$$

In particular all the directional derivatives exist at $(x, y) = 0$. However, f is not differentiable at 0. There are different ways to see this

- Note that $D_u f(0)$ is not linear in u!!! This is bad, since we showed that $D_u f(0) = Df(0) \cdot u$ provided that f is differentiable, and $Df(0)$ is a linear transformation. Thus f is not differentiable.

- Note that f is not even continuous at 0. If we approach $(0, 0)$ along the parabola $y = x^2$ we get that:

$$f(x, x^2) = \frac{x^4}{x^4 + x^4} = \frac{1}{2} \to 0$$

As $x \to 0$, but $f(0,0) = 0$.

The matrix whose entries are $\frac{\partial f_i}{\partial x_j}$ is called the Jacobian matrix. What we have learned up until now is:

- If f is differentiable at x then $Df(x)$ is equal to the Jacobian matrix at x.
- But the Jacobian matrix can exist without the derivative existing

Continuously differentiable functions

At this point he only criterion of differentiability at x that we can use is to go back to the definition. However, given how easy it is to compute partial derivative, it would be useful to have a criterion of differentiability based on partial derivatives.

Theorem. *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ where A is open. Suppose that all partial derivatives exist $\frac{\partial f}{\partial x_j} : U \to \mathbb{R}^m$ exist in some neighborhood U of $x \in A$ and they are all continuous at x. Then f is differentiable at x.*

In particular if all partial derivatives exist and are continuous through A, then f is differentiable in A. We call such an f a continuously differentiable, or C^1, function. This implies that $Df : \mathbb{R}^n \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ is continuous as well (since each of its component functions are continuous).

Proof. Next time!

Handout 4

- **Wish list for a measure function** It would be grand to have a measure function that tells us how big or small a subset of $\mathbb{R}^d$ is. This would be a function from the set of subsets of $\mathbb{R}^d$ into $[0, \infty]$, say $m : \mathcal{P}(\mathbb{R}^d) \to [0, \infty]$. We would like this function to satisfy the following properties:

  a) If $E_1, E_2, \ldots$ is a countable collection of disjoint subsets of $\mathbb{R}$, then
  $$m(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} m(E_n).$$
  This is called **Countable Additivity.**

  b) If E is congruent to F (i.e. F can be obtained from E by applying rigid motions: translations, rotations, or a reflections) then we should have that $m(E) = m(F)$.

  c) $m([0, 1)^d) = 1$.

  The bad news is that no such function can exist, and here's why (at least when $d = 1$). Let us define an equivalence relation between elements of $[0, 1)$ as follows: We say $x \sim y$ if $x - y$ is a rational number. Let N be the subset of $[0, 1)$ that contains exactly one element of each equivalence relation (the existence of this N requires invoking the axiom of choice). Now let $R = [0, 1) \cap \mathbb{Q}$, and for each $r \in R$ define the set

  $$N_r = \{x + r : x \in N \cap [0, 1 - r)\} \cup \{x + r - 1 : x \in N \cap [1 - r, 1)\}.$$

  (Basically N_r is just the translate of N by r units to the right, except that we move the part that sticks out of the interval $[0, 1)$ one unit to the left).

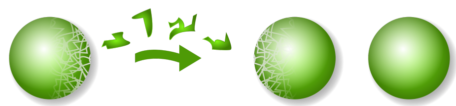


Figure 1: Banach-Tarski tells us that we can split the unit ball in $\mathbb{R}^3$ into finitely many (actually 5 is sufficient) many disjoint pieces, apply rigid motions to those pieces and then reassemble them to obtain two copies of the unit ball.

Q1) Show that $[0, 1)$ is the disjoint union of N_r for $r \in R$.

Q2) Show that if a measure function satisfying a), b) and c) above exists, then $m(N) = m(N_r)$ for every $r \in R$.

Q3) Arrive at a contradiction.

Remark: One might think that possibly relaxing condition a) to cover only *finitely* many disjoint sets E_n, i.e.

$$m(\cup_{n=1}^{N} E_n) = \sum_{n=1}^{N} m(E_n). \qquad \textbf{(Finite Additivity)}$$

would resolve the contradiction. Unfortunately, the Banach-Tarski paradox (cf. Figure 1) tells us that this is not enough to resolve this issue.

Conclusion: The problem with the above wishlist is that we insisted on being able to measure *every* subset of $\mathbb{R}^d$. We have shown that this is impossible. The solution is to be content with a measure function that is defined on some but not all subsets. Such subsets will be called measurable subsets.

The Greek method

- **Elementary measure.** An interval I is a subset of $\mathbb{R}$ of the form $[a, b], [a, b), (a, b], or (a, b)$ where $a, b \in \mathbb{R}$. The length of I is defined to be $|I| := b - a$. A *box* in $\mathbb{R}^d$ is a Cartesian product of intervals $B = I_1 \times I_2 \times \ldots I_d$ and its *volume* is defined to be $|B| = |I_1|. \ldots .|I_d|$. An *elementary set* is any subset of $\mathbb{R}^d$ which is the union of a finite number of boxes.

Q4) Show that if $E, F \subset \mathbb{R}^d$ are elementary sets, then the union $E \cup F$, the intersection $E \cap F$, the set theoretic difference $E \setminus F$, and the symmetric difference $E \Delta F = (E \setminus F) \cup (F \setminus E)$ are also elementary. Also, if $x \in \mathbb{R}^d$, then the translate $E + x := \{y + x : y \in E\}$ is also elementary.

Q5) Show that E can be expressed as the finite union of disjoint boxes. *Hint: Start with $d = 1$. Then use this result to generalize it to higher dimensions.*

- **Definition.** Let E be an elementary set. The above question allows to write $E = B_1 \cup B_2 \cup \ldots B_n$ where $B_1, \ldots, B_n$ are disjoint. We define the elementary measure of E as $m(E) := |B_1| + |B_2| + \ldots + |B_n|$.

Q6) Show that $m(E)$ is well-defined in the sense that if E can be expressed in two ways as a union of disjoint boxes $B_1, \ldots B_n$ and $B'_1, \ldots B'_m$, then

$$|B_1| + |B_2| + \ldots + |B_n| = |B'_1| + |B'_2| + \ldots + |B'_m|.$$

Hint: There's more than one approach you can take. One is to notice that for an interval I in $\mathbb{R}$, there holds that

$$|I| = \lim_{N \to \infty} \frac{1}{N} \# \left(I \cap \frac{1}{N} \mathbb{Z} \right).$$

(why?). And more generally for a box B,

$$|B| = \lim_{N \to \infty} \frac{1}{N^d} \# \left(B \cap \frac{1}{N} \mathbb{Z}^d \right).$$

Here $\frac{1}{N}\mathbb{Z}^d = \{\frac{k}{N} : k \in \mathbb{Z}^d\}$. Use this to give an alternative definition of $m(E)$ for an elementary set that does rely on its decomposition into disjoint boxes .

MATH 395 Notes

Faye Jackson

September 25, 2020

The wish list:

a) Say that $E_1, E_2, E_3, \ldots$ are disjoint subsets of $\mathbb{R}^d$, then:

$$m\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} m(E_n)$$

This is Countable Additivity

b) If E is congruent to F via translations, rotations, and combinations of these. We want to have $m(E) = m(F)$.

c) We want $m\left([0,1)^d\right) = m\left([0,1]^d\right) = 1$.

We know from last week and Q1-Q3 that there cannot be an $m : P(\mathbb{R}^d) \to [0, \infty]$ that satisfies a), b), and c). We construct the set $N \subseteq [0,1)$ containing exactly one element of each equivalence class for $x \sim y$ defined by $x - y \in \mathbb{Q}$. We define N_r as "essentially translates" of N by $r \in \mathbb{Q} \cap [0,1)$. Then in fact:

$$[0,1) = \coprod_{r \in \mathbb{Q} \cap [0,1)} N_r$$

And this union is disjoint. Furthermore $m(N_r) = N$ for each r because of congruence, so:

$$1 = m([0,1)) = \sum_{r \in \mathbb{Q} \cap [0,1)} m(N_r) = \sum_{n=1}^{\infty} m(N)$$

And whatever we choose for the measure of N, this produces a contradiction.

Proof of Q4. Let's go!

- First note that if $E = \bigcup_{i=1}^n A_i$ and $F = \bigcup_{k=1}^m B_k$ for some boxes A_i and B_k, then set $C_i = A_i$ if $1 \leq i \leq n$ and $C_i = B_{i-n}$ if $n < i \leq m+n$:

$$E \cup F = \bigcup_{i=1}^{n} A_i \cup \bigcup_{k=1}^{m} B_k = \bigcup_{i=1}^{m+n} C_i$$

  And so we have that $E \cup F$ is an elementary set as desired.

- We wish to show that $E \cap F$ is an elementary set for elementary sets E and F, **TODO**

- We wish to show that $E \setminus F$ is an elementary set for elementary sets E and F, **TODO**

- Note now that for elementary sets E and F we know:

$$E \triangle F = (E \setminus F) \cup (F \setminus E)$$

  And so by the previous bullets $E \setminus F$ and $F \setminus E$ are elementary, and so their union $E \triangle F$ is elementary

- We wish to show that the translate $E + x$ is elementary for an elementary E. **TODO**

With this we win!

Proof of Q5. induct **TODO**

Proof of Q6. **TODO**

MATH 395 Notes

Faye Jackson

September 28, 2020

Continuously Differentiable Functions

We saw that if $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at x then $Df(x)$ is given by the partial derivatives as:

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

The converse statement that if the partial derivatives $\frac{\partial f_i}{\partial x_j}$ exist for each $1 \leq i \leq m$ and $1 \leq j \leq n$ then Df exists is FALSE. However we have a slightly stronger condition that works!

Theorem. *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ where A is open. Suppose that all partial derivatives $\frac{\partial f_i}{\partial x_j}$ $(1 \leq i \leq m, 1 \leq j \leq n)$ exist in some neighborhood of $x \in A$, and they are continuous at x.*

Then f is differentiable at x. In particular if all partial derivatives exist and are continuous throughout A then f is differentiable in A. Such a function is called a C^1 function.

Remark. This theorem allows us to recognize "most" differentiable functions that we meet in practice just by checking that the partials are continuous.

Proof. Since f is differentiable at x if and only if each of its component functions are differentiable at x, we may assume without loss of generality that $m = 1$.

Let $r > 0$ be such that $B(x, r) \subseteq A$ and the partials are defined and continuous on $B(x,r)$. Then let $h \in \mathbb{R}^n$ such that $\|h\| < r$. Let $h = (h_1, \ldots, h_n)$. Set:

$$\begin{aligned} p_0 &:= x \\ p_k &:= p_{k-1} + h_k e_k \end{aligned}$$

And so $p_n = x + h$. So then we have:

$$f(x+h) - f(x) = \sum_{j=1}^{n} f(p_j) - f(p_{j-1})$$

Now we know that:

$$f(p_j) - f(p_{j-1}) = f(p_{j-1} + h_j e_j) - f(p_j)$$

Define $\phi_j(s) := f(p_{j-1} + s e_j)$ where ϕ is defined on some neighborhood of 0 in $\mathbb{R}$. Since ϕ_j is differentiable on an open interval containing $[0, h_j]$ with derivative $\frac{\partial f}{\partial x_j}$, we know that ϕ_j is continuous on $[0, h_j]$ and differentiable on $(0, h_j)$. Therefore by the mean value theorem we know that for some $c_j^h \in (0, h_j)$ that:

$$\begin{aligned} \phi_j(h_j) - \phi(0) &= h_j \cdot \phi'(c_j^h) \\ f(p_j) - f(p_{j-1}) &= h_j \cdot \frac{\partial f}{\partial x_j}(p_{j-1} + c_j^h e_j) \\ &= h_j \cdot \frac{\partial f}{\partial x_j}(q_j) \end{aligned}$$

Where q_j is some point in $B(x, \|h\|)$. Therefore:

$$f(x+h) - f(x) = \sum_{i=1}^{n} h_j \cdot \frac{\partial f}{\partial x_j}(q_j)$$

For some $q_1, \ldots, q_n$ in the ball of radius $\|h\|$ centered at x. Therefore:

$$f(x+h) - f(x) - \sum_{i=1}^{n} h_j \cdot \frac{\partial f}{\partial x_j}(x) = \sum_{j=1}^{n} h_j \cdot \left[\frac{\partial f}{\partial x_j}(q_j) - \frac{\partial f}{\partial x_j}(x) \right]$$

This implies that:

$$\frac{\left\| f(x+h) - f(x) - \sum_{i=1}^{n} h_j \cdot \frac{\partial f}{\partial x_j}(x) \right\|}{\|h\|} = \frac{\left\| \sum_{j=1}^{n} h_j \cdot \left[\frac{\partial f}{\partial x_j}(q_j) - \frac{\partial f}{\partial x_j}(x) \right] \right\|}{\|h\|}$$
$$\stackrel{\triangle}{\leq} \sum_{j=1}^{n} \frac{\|h_j\|}{\|h\|} \cdot \left\| \frac{\partial f}{\partial x_j}(q_j) - \frac{\partial f}{\partial x_j}(x) \right\|$$
$$\leq \sum_{j=1}^{n} \left\| \frac{\partial f}{\partial x_j}(q_j) - \frac{\partial f}{\partial x_j}(x) \right\|$$
$$\leq n \cdot \sup_{\substack{q \in B(x, \|h\|) \\ 1 \leq j \leq n}} \left\| \frac{\partial f}{\partial x_j}(q) - \frac{\partial f}{\partial x_j}(x) \right\|$$

But this goes to 0 as $\|h\| \to 0$ since $\frac{\partial f}{\partial x_j}$ are assumed to be continuous at x. Note then that we win! The function:

$$T(h) = \sum_{i=1}^{n} h_j \cdot \frac{\partial f}{\partial x_j}(x)$$

is a linear function, and so f is differentiable at x, and of course Df is just the vector $\left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)(x)$. Amazing!!!

Higher Order Derivatives

Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ where A is open. The component functions are $f_i : A \subseteq \mathbb{R}^n \to \mathbb{R}$. Since $\frac{\partial f_i}{\partial x_j}$ is itself a function from $A \to \mathbb{R}$ we can take higher order partial derivatives. For instance, if $f_i \in C^1$ then $\frac{\partial f_i}{\partial x_j}$ is defined and continuous, so we can consider if the following exists:

$$\frac{\partial^2 f_i}{\partial x_k \partial x_j} := \frac{\partial}{\partial x_k} \frac{\partial f_i}{\partial x_j}$$

This is called a second-order partial derivative. Similarly one can define partial derivatives of higher order inductively.

Definition. *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is of class $\underline{C^r}$ for $r \in \mathbb{N}_0$ provided that all the partial derivatives of its component functions up to order r exist and are continuous.*

We say that f is $\underline{C^\infty}$ provided that $f \in C^r$ for all $r \in \mathbb{N}_0$.

Exercise. *Show that $f \in C^r$ if and only if $\frac{\partial f}{\partial x_j} \in C^{r-1}$ for each $1 \leq j \leq n$.*

Multi-Index Notation

Definition. *A multi-index is an n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ such that each $\alpha_i \in \mathbb{N}_0$. If α is a multi-index then we define:*

- *The order of α as $|\alpha| := \alpha_1 + \cdots \alpha_n$. And the*
- *The factorial $\alpha! := \alpha_1!\alpha_2! \cdots \alpha_n!$*
- *For $x \in \mathbb{R}^n$ we define $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$.*
- *For $f_i : \mathbb{R}^n \to \mathbb{R}$ we define $\partial^\alpha f := \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} f$.*

Example. For $f : \mathbb{R}^2 \to \mathbb{R}$ we see:

$$\partial^{(1,2)} f = \frac{\partial^3 f}{\partial x_1 \partial x_2^2}$$

But wait, then what about $\frac{\partial^2 f}{\partial x_2 \partial x_1}$? Does it have a multi-index notation?

Theorem. *Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \to \mathbb{R}$ be a function of class C^2. Then for each $x \in A$ we have:*

$$\frac{\partial^2}{\partial x_k \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_k}(x)$$

Corrolary. *If $f : A \to \mathbb{R}$ is of class C^r then for any $2 \leq m \leq r$ then:*

$$\frac{\partial^m f}{\partial x_{j_1} \cdots \partial x_{j_m}} = \frac{\partial^m f}{\partial x_{\widetilde{j}_1} \cdots \partial x_{\widetilde{j}_m}}$$

for any permutation $\widetilde{j}_1, \ldots, \widetilde{j}_m$ of $j_1, \ldots, j_m$. In particular we can always rearrange $j_1, \ldots, j_m$ such that $\widetilde{j}_1 \leq \cdots \leq \widetilde{j}_m$ and in that case there is a multi-index notation:

$$\frac{\partial^m f}{\partial x_{j_1} \cdots \partial x_{j_m}} = \frac{\partial^m f}{\partial x_{\widetilde{j}_1} \cdots \partial x_{\widetilde{j}_m}} = \partial^\alpha f$$

For some multi-index α. Therefore any partial derivative up to order r can be written in multi-index notation as $\partial^\alpha f$ for some multi-index with order less than or equal to r.

Exercise. *Deduce the corollary from the theorem using induction.*

Proof of Theorem. We start with some reductions. Since one computes $\frac{\partial f}{\partial x_i \partial x_j}$ and $\frac{\partial f}{\partial x_j \partial x_i}$ by keeping all other coordinates x_k for $k \neq i, j$ constant, we can assume without loss of generality that $n = 2$, and that $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$.

MATH 395 Notes

Faye Jackson

September 30, 2020

Continue proving the equality of mixed partials

Theorem. *If $f \in C^2(A)$ where $A \subseteq \mathbb{R}^d$ then for each $x_0 \in A$ we have:*

$$\frac{\partial^2 f}{\partial x_k \partial x_j}(x_0) = \frac{\partial^2 f}{\partial x_j \partial x_k}(x_0)$$

Corrolary. *Equality of mixed partials of order r when $f \in C^r(A)$.*

Proof. We began by reducing to the case where $d = 2$, since in general all variables different from k, j are frozen when taking these partial derivatives. Thus assume $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ is C^2. Instead of referring to x_1, x_2 we'll refer to x, y.

Now lets consider our intuition. We know that $\frac{\partial f}{\partial x}$ measures $\frac{\Delta_x f}{h} = \frac{f(x_0+h,y)-f(x_0,y)}{h}$. And then:

$$\begin{aligned}
\frac{\partial^2 f}{\partial y \partial x} \approx \frac{\Delta_y \Delta_x f}{hk} &= \frac{\Delta_x f(x, y+h) - \Delta_x f(x,y)}{hk} \\
&= \frac{1}{hk}\left[f(x+h, y+k) - f(x, y+k) - f(x+h, y) + f(x,y)\right]
\end{aligned}$$

Similarly:

$$\begin{aligned}
\frac{\partial f}{\partial y} \approx \frac{\Delta_y f}{k} &= \frac{f(x, y+k) - f(x,y)}{k} \\
\frac{\partial^2 f}{\partial x \partial y} \approx \frac{\Delta_x \Delta_y f}{hk} &= \frac{\Delta_y f(x+k, y) - \Delta_y f(x,y)}{hk} \\
&= \frac{1}{hk}\left[f(x+h, y+k) - f(x+h, y) - f(x, y+k) + f(x,y)\right]
\end{aligned}$$

Notice that $\Delta_y \Delta_x f = \Delta_x \Delta_y f$. Thus the equality of this discrete version of the partials that we expect the partials to be the same.

Now for the real proof. Let $(x_0, y_0) \in A$ and Q be the rectangle with vertices (x_0, y_0), $(x_0 + h, y_0)$, $(x_0, y_0 + k)$, $(x_0 + h, y + k)$. Since A is open, we can take h and k to be small enough so that $Q \subseteq A$. Now let:

$$G(h, k) = f(x_0 + h, y_0 + k) - f(x_0 + k, y_0) - f(x_0, y_0 + k) + f(x_0, y_0)$$

We will show that:

$$G(h, k) = hk\frac{\partial^2 f}{\partial x \partial y}(p) = hk\frac{\partial^2 f}{\partial y \partial x}(q)$$

For some $p, q \in Q$. To show the first equality. Let us use $G(h, k) = \Delta_y \Delta_x f$ and let $\phi(y) = f(x_0 + h, y) - f(x_0, y)$ for y between y_0 and $y_0 + k$. We know that ϕ is continuous on $[y_0, y_0 + k]$ since f itself is. Also ϕ is differentiable on $(y_0, y_0 + k)$ since f is C^1. Therefore by the Mean Value Theorem there exists a $y_\star$ between y_0 and $y_0 + k$ so that:

$$\phi(y_0 + k) - \phi(y_0) = \phi'(y_\star)k$$

Notice then that:

$$\begin{aligned} G(h, k) &= \phi(y_0 + k) - \phi(y_0) \\ \phi'(y) &= \frac{\partial f}{\partial y}(x_0 + h, y) - \frac{\partial f}{\partial y}(x_0, y) \\ G(h, k) &= k\left[\frac{\partial f}{\partial y}(x_0 + h, y_\star) - \frac{\partial f}{\partial y}(x_0, y_\star)\right] \end{aligned}$$

Now we know that $\frac{\partial f}{\partial y}(x, y_\star)$ is continuous on the closed interval between x_0 and $x_0 + h$ and differentiable on the open interval. By the MVT there is a $x_\star$ between x_0 and $x_0 + h$ so that:

$$G(h, k) = kh\frac{\partial^2 f}{\partial x \partial y}(x_\star, y_\star) = k\left[\frac{\partial f}{\partial y}(x_0 + h, y_\star) - \frac{\partial f}{\partial y}(x_0, y_\star)\right]$$

Note that $(x_\star, y_\star) \in Q$ so we have the first equality. To show the other equality, we argue similarly using the fact that $G(h, k) = \Delta_x \Delta_y f$. More precisely instead of ϕ above we introduce:

$$\psi(x) = f(x, y_0 + k) - f(x, y_0)$$

By MVT we can get a $x_\heartsuit$ such that:

$$G(h,k) = \psi(x_0 + h) - \psi(x_0) = h\psi'(x_\heartsuit)$$
$$G(h,k) = h\left[\frac{\partial f}{\partial x}(x_\heartsuit, y_0 + k) - \frac{\partial f}{\partial x}(x_\heartsuit, y_0)\right]$$

By applying the mean value theorem again we get $y_\heartsuit$ between y_0 and $y_0 + k$ we get:

$$G(h,k) = hk\frac{\partial^2 f}{\partial y \partial x}(x_\heartsuit, y_\heartsuit)$$

This is exactly the same moves as in the proof for x. Then:

$$\frac{G(h,k)}{hk} = \frac{\partial^2 f}{\partial x \partial y}(x_\star, y_\star) = \frac{\partial^2}{\partial y \partial x}(x_\heartsuit, y_\heartsuit)$$

By letting $h, k \to 0$ both $(x_\star, y_\star) = p \to (x_0, y_0)$ and $(x_\heartsuit, y_\heartsuit) = q \to (x_0, y_0)$. By continuity of $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ at (x_0, y_0) we obtain the desired equality that:

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$$

The Chain Rule and Taylor's Formula in Higher Dimensions

Recall. For $f : A \subseteq \mathbb{R} \to \mathbb{R}$ and $g : B \subseteq \mathbb{R} \to \mathbb{R}$ such that $f(A) \subseteq B$ we have $g \circ f : A \subseteq \mathbb{R} \to \mathbb{R}$. We have:

$$\frac{\mathrm{d}}{\mathrm{d}x}[g \circ f](x) = g'(f(x)) \cdot f'(x)$$

provided that $f'(x)$ and $g'(f(x))$.

Theorem (Chain Rule)**.** *Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ and suppose that $f : A \subseteq \mathbb{R}^m$ and $g : B \subseteq \mathbb{R}^k$. with $f(A) \subseteq B$. Suppose that x_0 is an interior point of A and $y_0 = f(x_0)$ is an interior point of B. Furthermore suppose that f is differentiable at x_0 and g is differentiable at y_0. Then $g \circ f$ is differentiable at x_0 and:*

$$D[g \circ f](x_0) = Dg(y_0) \circ Df(x_0) = Dg(f(x_0)) \cdot Df(x_0)$$

Proof. Since y_0 is an interior point of B there exists a $\varepsilon > 0$ such that $B(y_0, \varepsilon) \subseteq B$. Since f is continuous at x_0 there exists a $\delta > 0$ so that $f(B(x_0, \delta)) \subseteq B(y_0, \varepsilon)$. So we can define $g \circ f : B(x_0, \delta \to \mathbb{R}^k$. Let $\|h\| < \delta$ for $h \in \mathbb{R}^n$ and define:

$$R_f(h) = \frac{f(x_0 + h) - f(x_0) - Df(x_0) \cdot h}{\|h\|} \qquad (h \neq 0)$$

$$R_f(h) = 0 \qquad (h = 0)$$

By differentiability of f at x_0 we have $R_f(h) \to 0$ as $\|h\| \to 0$. Similarly if $\|k\| < \varepsilon$ and $k \in \mathbb{R}^m$ we define:

$$R_g(k) = \frac{g(y_0 + k) - g(y_0) - Dg(y_0) \cdot k}{\|k\|} \qquad (k \neq 0)$$

$$R_g(k) = 0 \qquad (k = 0)$$

By differentiability we know that $R_g(k) \to 0$ as $\|k\| \to 0$. To show that $g \circ f$ is differentiable at x_0 we must show that there exists an $A \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^k)$ such that:

$$R_{g \circ f}(h) = \frac{[g \circ f](x_0 + h) - [g \circ f](x_0) - Ah}{\|h\|} \to 0 \text{ as } \|h\| \to 0$$

Rewrite as the following:

$$\begin{aligned}
[g\circ f](x_0+h)-[g\circ f](x_0) &= g(f(x_0+h))-g(f(x_0))\\
&= g(f(x_0)+f(x_0+h)-f(x_0))-g(f(x_0))\\
&= g(y_0+k)-g(y_0)
\end{aligned}$$

Where we call $k=f(x_0+h)-f(x_0)$. From $R_g(k)$ we know that for any $k\in\mathbb{R}^m$:

$$g(y_0+k)-g(y_0)=Dg(y_0)\cdot k+\|k\|R_g(k)$$

Furthermore $k=f(x_0+h)-f(x_0)=Df(x_0)+\|h\|R_f(h)$. Therefore:

$$\begin{aligned}
g(y_0+k)-g(y_0) &= Dg(y_0)[Df(x_0)h+\|h\|R_f(h)]+\|k\|R_g(k)\\
&= Dg(y_0)Df(x_0)\cdot h+\|h\|D_g(y_0)R_f(h)+\|k\|R_g(k)
\end{aligned}$$

Set $A=Dg(y_0)\cdot Df(x_0)$ This gives that for $h\neq 0$ that:

$$\begin{aligned}
R_{g\circ f}(h) &= \frac{[g\circ f](x_0+h)-[g\circ f](x_0)-Ah}{\|h\|}\\
&= Dg(y_0)R_f(h)+\frac{\|k\|}{\|h\|}R_g(k)
\end{aligned}$$

We know that $R_f(h)\to 0$ as $\|h\|\to 0$. It remains to show that $\frac{\|k\|}{\|h\|}R_g(k)\to 0$ as $\|h\|\to 0$. We know that:

$$\begin{aligned}
\|k\| &= \|Df(x_0)\cdot h+\|h\|R_f(h)\|\\
&\leq \|Df(x_0)\cdot h\|+\|h\|\|R_f(h)\|\\
&\leq C\|h\|+\|h\|\leq (C+1)\|h\|
\end{aligned}$$

This follows since $\|R_f(h)\|\leq 1$ if $\|h\|$ is small enough. Also we kow since $Df(x_0)$ is linear we know $\|Df(x_0)\cdot h\|\leq C\|h\|$ for some constant C by 296 / linear algebra. Therefore:

$$\left\|\frac{\|k\|}{\|h\|}R_g(k)\right\|\leq (C+1)\frac{\|h\|}{\|h\|}\|R_g(k)\|\leq (C+1)\|R_g(k)\|$$

Therefore as $\|h\|\to 0$ we know that $\|k\|\to 0$ since $\|k\|\leq (C+1)\|h\|$ and hence $R_g(k)\to 0$. Therefore $\frac{\|k\|}{\|h\|}R_g(k)\to 0$ as $h\to 0$ and so this finishes the proof.

Taylor's Theorem in several variables

Recall the multi-index notation from last time.

Lemma (The multinomial lemma)**.** *For any $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and any positive integer k we have:*

$$(x_1 + x_2 + \cdots + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha$$

Handout 5

The Elementary measure (Continued)

- Recall from last time that an interval I is a subset of $\mathbb{R}$ of the form $[a,b],[a,b),(a,b],or(a,b)$ where $a,b\in\mathbb{R}$. The length of I is defined to be $|I| := b-a$. A *box* in $\mathbb{R}^d$ is a Cartesian product of intervals $B = I_1 \times I_2 \times \dots I_d$ and its *volume* is defined to be $|B| = |I_1|.\dots.|I_d|$. An *elementary set* is any subset of $\mathbb{R}^d$ which is the union of a finite number of boxes.

- **Definition.** Let E be an elementary set. Last time we saw that we can write $E = B_1 \cup B_2 \cup \dots B_n$ where $B_1,\dots,B_n$ are disjoint. We define the elementary measure of E as $m(E) := |B_1| + |B_2| + \dots + |B_n|$.

Q1) Show that $m(E)$ is well-defined in the sense that if E can be expressed in two ways as a union of disjoint boxes $B_1,\dots B_n$ and $B_1',\dots B_m'$, then

$$|B_1| + |B_2| + \dots + |B_n| = |B_1'| + |B_2'| + \dots + |B_m'|.$$

Hint: There's more than one approach you can take. One is to notice that for an interval I in $\mathbb{R}$, there holds that

$$|I| = \lim_{N\to\infty} \frac{1}{N}\#\left(I \cap \frac{1}{N}\mathbb{Z}\right).$$

(why?). And more generally for a box B,

$$|B| = \lim_{N\to\infty} \frac{1}{N^d}\#\left(B \cap \frac{1}{N}\mathbb{Z}^d\right).$$

Here $\frac{1}{N}\mathbb{Z}^d = \{\frac{k}{N} : k \in \mathbb{Z}^d\}$. Use this to give an alternative definition of $m(E)$ for an elementary set that does rely on its decomposition into disjoint boxes.

- **Properties of Elementary measure.** Show that the following holds

Q2) Show that if $E_1, \ldots, E_n$ are disjoint elementary sets, then

$$m(E_1 \cup \ldots \cup E_n) = \sum_{i=1}^{n} m(E_i)$$

Recall that this is called **finite additivity**.

Q3) Show that if $E \subset F$ are two elementary sets, then

$$m(E) \leq m(F).$$

This property is called **monotonicity**.

Q4) Show that if $E_1, E_2, \ldots, E_n$ is an arbitrary finite collection of elementary sets, then

$$m(E_1 \cup \ldots E_n) \leq m(E_1) + \ldots + m(E_n).$$

This is called **finite subadditivity**.

- **Why is this unsatisfactory?** Of course, the main problem with this measure is that we can only measure relatively simple sets (namely the elementary sets). For example, we cannot measure the area of a disc. One might be tempted to generalize this measure naively as follows: For an arbitrary set $E \subset \mathbb{R}^d$, define

$$m_{\text{pixel}}(E) = \lim_{N \to \infty} \frac{1}{N^d} \# \left(E \cap \frac{1}{N}\mathbb{Z}^d \right).$$

However, this is not a particurlary satisfactory definition for (at least) the following two reasons:

Q5) Find a subset E of $\mathbb{R}$ for which this limit does not exist.

Q6) Find a subset E of $\mathbb{R}$ such that both $m_{\text{pixel}}(E)$ and $m_{\text{pixel}}(E+x)$ exist, but $m_{\text{pixel}}(E) \neq m_{\text{pixel}}(E+x)$ for some $x \in \mathbb{R}$.

MATH 395 Notes

Faye Jackson

October 2, 2020

Proof of Q1.

Lemma. *For any interval I in $\mathbb{R}$ we have that:*

$$|I| = \lim_{N\to\infty} \frac{1}{N} \cdot \# \left(I \cap \frac{1}{N}\mathbb{Z} \right)$$

Proof. Consider that the following sets are in bijection:

$$\begin{aligned} f : I \cap \frac{1}{N}\mathbb{Z} &\to NI \cap \mathbb{Z} \\ x &\mapsto N \cdot I \end{aligned}$$

This maps its domain into the codomain by definition, since $N \cdot I = \{N \cdot x \mid x \in I\}$ and $\frac{1}{N}\mathbb{Z} = \left\{\frac{1}{N} \cdot m \mid m \in \mathbb{Z}\right\}$. We also know since $N > 0$ that this is an injection from linear algebra. We also know surjectivity as well by quick definition from the sets. Now say I has endpoints $a \leq b$, then NI has endpoints aN and bN.

Now note that the cardinality $\#(NI \cap \mathbb{Z})$ is between $bN - aN - 5$ and $bN - aN + 5$. So then note that:

$$\begin{aligned} bN - aN - 5 \leq \#(NI \cap \mathbb{Z}) &\leq bN - aN + 5 \\ b - a - \frac{5}{N} \leq \frac{1}{N}\#\left(I \cap \frac{1}{N}\mathbb{Z}\right) &\leq b - a + \frac{5}{N} \\ b - a \leq \lim_{N\to\infty} \frac{1}{N}\#\left(I \cap \frac{1}{N}\mathbb{Z}\right) &\leq b - a \end{aligned}$$

By squeeze theorem! We win! This limit is equal to $|I| = b - a$.

Lemma. *For any box $B \subseteq \mathbb{R}^d$, we have:*

$$|B| = \lim_{N\to\infty} \frac{1}{N^d} \# \left(B \cap \frac{1}{N}\mathbb{Z}^d \right)$$

Proof. First write $B = \prod_{k=1}^d I_k$ for intervals I_k and note that:

$$B \cap \frac{1}{N}\mathbb{Z}^d = \left(\prod_{k=1}^d I_k \right) \cap \prod_{k=1}^d \frac{1}{\mathbb{N}} \cdot \mathbb{Z} = \prod_{k=1}^d \left(I_k \cap \frac{1}{\mathbb{N}} \cdot \mathbb{Z} \right)$$

$$\# \left(B \cap \frac{1}{N}\mathbb{Z}^d \right) = \# \left(\prod_{k=1}^d \left(I_k \cap \frac{1}{\mathbb{N}} \cdot \mathbb{Z} \right) \right) = \prod_{k=1}^d \# \left(I_k \cap \frac{1}{\mathbb{N}} \cdot \mathbb{Z} \right)$$

So now we write that:

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N^d} \# \left(B \cap \frac{1}{N}\mathbb{Z}^d \right) &= \lim_{N\to\infty} \frac{1}{N^d} \prod_{k=1}^d \# \left(I_k \cap \frac{1}{\mathbb{N}} \cdot \mathbb{Z} \right) \\ &= \prod_{k=1}^d \lim_{N\to\infty} \frac{1}{N} \# \left(I_k \frac{1}{N}\mathbb{Z} \right) \\ &= \prod_{k=1}^d |I_k| = |B| \end{aligned}$$

And therefore the lemma is proved!

We prove one final lemma, and then the result will fall out!

Lemma. *Suppose that we have two disjoint sets $X, Y \subseteq \mathbb{R}^d$ and the limits:*

$$\lim_{N\to\infty} \frac{1}{N^d} \# \left(X \cap \frac{1}{N}\mathbb{Z}^d \right) \qquad \lim_{N\to\infty} \frac{1}{N^d} \# \left(Y \cap \frac{1}{N}\mathbb{Z}^d \right)$$

both exist, then:

$$\lim_{N\to\infty} \frac{1}{N^d} \# \left((X \cup Y) \cap \frac{1}{N}\mathbb{Z}^d \right) = \lim_{N\to\infty} \frac{1}{N^d} \# \left(X \cap \frac{1}{N}\mathbb{Z}^d \right) + \lim_{N\to\infty} \frac{1}{N^d} \# \left(Y \cap \frac{1}{N}\mathbb{Z}^d \right)$$

Proof. This is fairly simple to prove. Note that:

$$(X \cup Y) \cap \frac{1}{N}\mathbb{Z}^d = \left(X \cap \frac{1}{N}\mathbb{Z}^d\right) \cup \left(Y \cap \frac{1}{N}\mathbb{Z}^d\right)$$

And since these are disjoint:

$$\begin{aligned}
\#\left((X \cup Y) \cap \frac{1}{N}\mathbb{Z}^d\right) &= \#\left(\left(X \cap \frac{1}{N}\mathbb{Z}^d\right) \cup \left(Y \cap \frac{1}{N}\mathbb{Z}^d\right)\right) \\
&= \#\left(X \cap \frac{1}{N}\mathbb{Z}^d\right) + \#\left(Y \cap \frac{1}{N}\mathbb{Z}^d\right)
\end{aligned}$$

We then know that we can take the limit as $N \to \infty$ on either side by real analysis and we must get the same limit as desired in the lemma

Now fix an elementary set $E \subseteq \mathbb{R}^d$ and let it be the union of disjoint boxes $B_1, \ldots, B_n$. By applying the lemmas multiple times:

$$\begin{aligned}
m(E) = \sum_{k=1}^{n} |B_k| &= \sum_{k=1}^{n} \lim_{N\to\infty} \frac{1}{N^d} \# \left(B_k \cap \frac{1}{N}\mathbb{Z}^d\right) \\
&= \lim_{N\to\infty} \frac{1}{N^d} \# \left(\left(\bigcup_{k=1}^{n} B_k\right) \cap \frac{1}{N}\mathbb{Z}^d\right) \\
&= \lim_{N\to\infty} \frac{1}{N^d} \left(E \cap \frac{1}{N}\mathbb{Z}^d\right)
\end{aligned}$$

Now note that the limit does not depend on the choice of disjoint boxes $B_1, \ldots, B_n$, so if we choose another choice of disjoint boxes $B_1', \ldots, B_m'$ that union to E then we know:

$$\sum_{k=1}^{n} |B_k| = \lim_{N\to\infty} \frac{1}{N^d} \left(E \cap \frac{1}{N}\mathbb{Z}^d\right) = \sum_{k=1}^{m} |B_k'|$$

And so the measure of E is well-defined.

Taylor's Theorem on $\mathbb{R}^d$

Lemma (The multinomial lemma)**.** *Let $x = (x_1, \ldots, x_n)$. We would like to look at:*

$$(x_1 + \cdots + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha$$

With:

$$\begin{aligned} |\alpha| &= \alpha_1 + \cdots \alpha_n \\ \alpha! &= \alpha_1! \cdots \alpha_n! \\ x^\alpha &= x_1^{\alpha_1} \cdots x_n^{\alpha_n} \end{aligned}$$

This generalizes the binomial theorem.

Proof. The proof proceeds by induction on n. The binomial theorem gives the case $n = 2$. Suppose that the multinomial theorem holds up to $n-1$. We want to show it holds for n, where $n \geq 3$. So then we write:

$$\begin{aligned} (x_1 + x_2 + \cdots + x_n)^k = (x_1 + (x_2 + \cdots + x_n))^k &= \sum_{a+b=k} \frac{k!}{a!b!} x_1^a (x_2 + \cdots + x_n)^b \\ &= \sum_{a+b=k} \frac{k!}{a!b!} x_1^a \sum_{|\beta|=b} \frac{b!}{\beta!} (x_2, \ldots, x_n)^\beta \\ &= \sum_{a+b=k} \sum_{\substack{|\beta|=b \\ \beta \in \mathbb{N}_0^{n-1}}} \frac{k!}{a!\beta!} x_1^a x_2^{\beta_1} \cdots x_n^{\beta_{n-1}} \end{aligned}$$

Now set $\alpha = (a, \beta)$. Then:

$$\begin{aligned} (x_1 + x_2 + \cdots + x_n)^k &= \sum_{a+b=k} \sum_{\substack{|\beta|=b \\ \beta \in \mathbb{N}_0^{n-1}}} \frac{k!}{a!\beta!} x_1^a x_2^{\beta_1} \cdots x_n^{\beta_{n-1}} \\ &= \sum_{\substack{|\alpha|=k \\ \alpha \in \mathbb{N}_0^n}} \frac{k!}{\alpha!} x^\alpha \end{aligned}$$

Therefore the result follows by induction. Great!!!

Lemma (Higher order product rule)**.** *For any $\alpha \in \mathbb{N}_0^n$ and $f, g : \mathbb{R}^n \to \mathbb{C}$ we have:*

$$\partial^\alpha (fg) = \sum_{\substack{\beta+\gamma=\alpha \\ \beta,\gamma\in\mathbb{N}_0^n}} \frac{\alpha!}{\beta!\gamma!} \partial^\beta f \partial^\gamma g$$

Whenever f and g are differentiable up to order $|\alpha|$. This generalizes Leibniz Rule.

Recall. We take as notation:

$$\partial_j^a f = \partial_{x_j}^a f = \frac{\partial^a f}{\partial x_j^a}$$

For convenience

Proof. Again the proof is by induction on n. For $n = 1$, let $\alpha = k \in \mathbb{N}_0$, we want to show that:

$$\partial^k (fg) = \sum_{p+q=k} \frac{k!}{p!q!} \partial^p f \partial^q g = \sum_{p=0}^{k} \frac{k!}{p!(k-p)!} \partial^p f \partial^{k-p} g$$

This is part of your homework. Press F to pay respects. THerefore the result is true when $n = 1$. Now assume the result is true for $n - 1$, we will show it holds for n.

Take $f, g : \mathbb{R}^n \to \mathbb{C}$ and take $\alpha \in \mathbb{N}_0^n$. Write $\alpha = (a, \theta)$ where $a \in \mathbb{N}_0$, $\theta \in \mathbb{N}_0^{n-1}$, and $x = (x_1, x')$ where $x_1 \in \mathbb{R}$ and $x' \in \mathbb{R}^{n-1}$. Then:

$$\begin{aligned}
\partial_x^\alpha (fg) = \partial_{x_1}^a \partial_{x'}^\theta (fg) &= \partial_{x_1}^\alpha \left[\sum_{\substack{\mu+\nu=\theta \\ \mu,\nu\in\mathbb{N}_0^{n-1}}} \frac{\theta!}{\mu!\nu!} \partial_{x'}^\mu f \partial_{x'}^\nu g \right] \\
&= \sum_{\substack{\mu+\nu=\theta \\ \mu,\nu\in\mathbb{N}_0^{n-1}}} \frac{\theta!}{\mu!\nu!} \partial_{x_1}^\alpha \left[\partial_{x'}^\mu f \partial_{x'}^\nu g \right] \\
&= \sum_{\substack{\mu+\nu=\theta \\ \mu,\nu\in\mathbb{N}_0^{n-1}}} \frac{\theta!}{\mu!\nu!} \sum_{m+k=a} \frac{a!}{m!k!} \partial_{x_1}^m \partial_{x'}^\mu f \partial_{x_1}^k \partial_{x'}^\nu
\end{aligned}$$

So then we may write:

$$
\begin{aligned}
\partial_x^\alpha(fg) &= \sum_{\substack{\mu+\nu=\theta \\ \mu,\nu\in\mathbb{N}_0^{n-1}}} \sum_{m+k=a} \frac{a!\theta!}{(\mu!m!)(\nu!k!)} \partial_{x_1}^m \partial_{x'}^\mu f \partial_{x_1}^k \partial_{x'}^\nu \\
&= \sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\gamma!} \partial^\beta f \partial^\gamma g
\end{aligned}
$$

The result now follows by induction. Great! Here we take:

Recall. We recall Taylor's Theorem for single-variable functions. Suppose $f : [a,b] \to \mathbb{R}$ is $C^k([a,b])$ and $\partial^k f : (a,b) \to \mathbb{R}$ is differentiable. Then for any $a \leq x \leq b$ then:

$$
\begin{aligned}
f(x) &= R_{a,k}(x) + \sum_{j=0}^{k} \frac{(x-a)^j \cdot f^{(j)}(a)}{j!} \\
R_{a,k}(x) &= \frac{(x-a)^{k+1}}{(k+1)!} f^{(k+1)}(c)
\end{aligned}
$$

For some $a \leq c \leq x$.

We will study the generalization of this theorem for functions $f : \mathbb{R}^n \to \mathbb{R}^m$.

Recall. If $f = (f_1, \dots, f_m)$ and α is a multi-index then:

$$
\partial^\alpha f = \begin{pmatrix} \partial^\alpha f_1 \\ \cdots \\ \partial^\alpha f_m \end{pmatrix}
$$

Thus we only need to consider the case $m = 1$

Definition. *We call a subset $G \subseteq \mathbb{R}^n$ convex provided that for every $x, y \in G$ and every $t \in [0,1]$ we have $tx + (1-t)y \in G$.*

The Plan: We would like to derive the Taylor Expansion of f at some point a of its domain (which should be open and convex). At order k this should give us a polynomial in $x_1, \dots, x_n$ of degree $\leq k$ that approximates the function near a.

The General Statement and Proof

Theorem (Taylor's Theorem)**.** *Let $G \subseteq \mathbb{R}^n$ be an open convex set. Suppose that $f : G \to \mathbb{C}$ is of class C^{k+1}. If $a \in G$, then for any $x \in G$ we have:*

$$f(x) = R_{a,k}(x) + \sum_{\substack{|\alpha| \leq k \\ \alpha \in \mathbb{N}_0^n}} \frac{1}{\alpha!}(x-a)^\alpha \partial^\alpha f(a)$$

where we have:

$$R_{a,k}(x) = \sum_{\substack{|\alpha| = k+1 \\ \alpha \in \mathbb{N}_0^n}} \frac{1}{\alpha!}(x-a)^\alpha \partial^\alpha f(c)$$

For some $c \in G$ on the line segment connecting a and x, that is $c = ta + (1-t)x$ for some $t \in [0,1]$.

Recall. Recall the following formula

$$D_u f(x+tu) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f(x+tu+su) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f(x+(t+s)u) = \frac{\mathrm{d}}{\mathrm{d}r}\bigg|_{r=t} f(x+ru)$$
$$D_u f(x+tu) = \frac{\mathrm{d}}{\mathrm{d}t} f(x+tu)$$

Which is nice

Proof. To avoid confusion, let us denote x by x_0. We will deduce this result from the single-variable case. To do so we will look at the restriction of f along the line segment connecting a and x_0, by convexity this line segment belongs to G. Set:

$$\phi : [0,1] \to \mathbb{C}$$
$$t \overset{\phi}{\mapsto} f(ta + (1-t)x_0)$$

Notice that $\phi(0) = f(a)$ and $\phi(1) = f(x_0)$, furthermore note that $\phi \in C^{k+1}([0,1])$ since $f \in C^{k+1}(G)$. By Taylor's Formula in one dimension at $t = 0$ we know:

$$\phi(1) = R_{0,k}(1) + \sum_{p=0}^{k} \frac{\phi^{(p)}(0) \cdot 1^p}{p!}$$
$$R_{0,k} = \frac{\phi^{k+1}(c)}{(k+1)!} \cdot 1^{k+1}$$

What is $\phi^{(p)}(0)$? For $p = 0$ we know $\phi^{(0)}(0) = \phi(0) = f(a)$. For $p = 1$ we have

$$\begin{aligned}
\phi^{(1)}(t) = \phi'(t) &= \frac{\mathrm{d}}{\mathrm{d}t} f(a + t(x_0 - a)) \\
&= Df(a + t(x_0 - a)) \cdot (x_0 - a) = D_u f(a + tu)
\end{aligned}$$

Where $u = x_0 - a$. But then this is equal to:

$$\phi'(t) = \left(u_1 \frac{\partial}{\partial x_1} + \cdots + u_n \frac{\partial}{\partial x_n} \right) f(a + tu)$$

So then we know that:

$$\phi'(0) = \left(u_1 \frac{\partial}{\partial x_1} + \cdots + u_n \frac{\partial}{\partial x_n} \right) f(a)$$

Now for $p = 2$:

$$\begin{aligned}
\phi''(t) &= \frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_{j=1}^{n} u_j \frac{\partial}{\partial x_j} \right) f(a + tu) \\
&= \left(\sum_{j=1}^{n} u_j \frac{\partial}{\partial x_j} \right) D_u f(a + tu) = \left(\sum_{j=1}^{n} u_j \frac{\partial}{\partial x_j} \right)^2 f(a + tu)
\end{aligned}$$

Think of these as operators on functions that we're manipulating and consider:

$$\frac{\mathrm{d}}{\mathrm{d}t} u_1 \frac{\partial f}{\partial x_1}(a + tu) = u_1 D_u \left(\frac{\partial f}{\partial x_1} \right) (x + tu)$$

And so in general we want to think about:

$$\begin{aligned}
\phi^{(p)}(t) &= \left(\sum_{j=1}^{n} u_j \frac{\partial}{\partial x_j} \right)^p f(a + tu) \\
\phi^{(p)}(0) &= \left(\sum_{j=1}^{n} u_j \frac{\partial}{\partial x_j} \right)^p f(a)
\end{aligned}$$

MATH 395 Notes

Faye Jackson

October 7, 2020

Theorem (Taylor's Theorem)**.** *Let G be open and convex. Let $f : G \to \mathbb{C}$ be C^{k+1} and $a \in G$. Then:*

$$f(x) = R_{a,k}(x) + \sum_{|\alpha| \le k} \frac{\partial^\alpha f(a)}{\alpha!}(x-a)^\alpha$$
$$R_{a,k}(x) = \sum_{|\alpha|=k+1} \frac{\partial^\alpha f}{\alpha!}(c)(x-a)^\alpha$$

Where c is on the line segment connecting a and x

Continued Proof of Taylor's Theorem. We'll fix some $x_0 \in G$. Then set $\phi(t) = f(a + t(x_0 - a))$ where $\phi : [0,1] \to \mathbb{C}$. By taylor's theorem in one dimension:

$$f(x_0) = \phi(1) = R_{0,k}(1) \sum_{p=0}^{k} \frac{\phi^{(p)}(0)}{p!} 1^p$$
$$R_{0,k}(1) = \frac{\phi^{(k+1)}(\theta)}{(k+1)!} 1^{k+1}$$

For some $0 \le \theta \le 1$. We need a formula for $\phi^{(p)}(t)$. Let $u = (x_0 - a)$. Then by the chain rule:

$$\phi'(t) = Df(a+tu) \cdot u = D_u f(a+tu) = \left(\sum_{k=1}^{n} u_k \frac{\partial f}{\partial x_k} \right)(a+tu)$$

So then if we call $f_1 = D_u f$ then we have that:

$$\begin{aligned}\phi''(t) &= D_u f_1(a+tu) = [D_u(D_u f)](a+tu)\\ &= D_u^2 f(a+tu) = \left(\sum_{k=1}^n u_k \frac{\partial}{\partial x_k}\right)^2 f(a+tu)\end{aligned}$$

So then by induction we can obtain that:

$$\phi^{(j)}(x) = D_u^j f(a+tu) = \left(\sum_{k=1}^n u_k \frac{\partial}{\partial x_j}\right)^j f(a+tu)$$

Where this holds for $0 \leq j \leq k+1$, since f is differentiable $k+1$ times. And so for $0 \leq j \leq k$ we have:

$$\begin{aligned}\phi^{(j)}(0) &= D_u^j f(a) = \left(\sum_{k=1}^n u_k \frac{\partial}{\partial x_k}\right)^j f(a)\\ \phi^{(k+1)}(\theta) &= D_u^{k+1} f(a) = \left(\sum_{k=1}^n u_k \frac{\partial}{\partial x_k}\right)^{k+1} f(a+\theta u)\end{aligned}$$

Consider that as operators we can show—using linearity—similarly to how we showed the multinomial lemma, we have:

$$\left(\sum_{k=1}^n u_k \frac{\partial}{\partial x_k}\right)^p = \sum_{|\alpha|=p} \frac{p!}{\alpha!} u^\alpha \partial^\alpha$$

This gives us that:

$$\begin{aligned}\phi^{(p)}(0) &= \sum_{|\alpha|=p} \frac{p!}{\alpha!} u^\alpha \partial^\alpha f(a)\\ \phi^{(k+1)}(\theta) &= \sum_{|\alpha|=k+1} \frac{(k+1)!}{\alpha!} u^\alpha \partial^\alpha f(a+\theta u)\end{aligned}$$

Set $c = a + \theta u$ which is on the line segment between a and x_0, so then we must have that:

$$
\begin{aligned}
f(x_0) = \phi(1) &= \frac{\phi^{(k+1)}}{(k+1)!} + \sum_{p=0}^{k} \frac{\phi^{(p)}(a)}{p!} \\
&= \frac{1}{(k+1)!} \sum_{|\alpha|=k+1} \frac{(k+1)!}{\alpha!} u^\alpha \partial^\alpha(c) \\
&\quad + \sum_{p=0}^{k} \frac{1}{p!} \left(p! \sum_{|\alpha|=p} \frac{1}{\alpha!} u^\alpha \partial^\alpha f(a) \right) \\
&= \sum_{|\alpha|\le k} \frac{\partial^\alpha f(a)}{\alpha!} u^\alpha + \sum_{|\alpha|=k+1} \frac{\partial^\alpha(c)}{\alpha!} u^\alpha
\end{aligned}
$$

This is exactly what we want to show!

Example. Let $f(x,y) = \sin(x^2 + y)$ where $f : \mathbb{R}^2 \to \mathbb{R}$. Find the 3rd degree polynomial that best approximates f near $(0,0)$.

This is simply:

$$
P(x,y) = \sum_{|\alpha|\le 3} \frac{\partial^\alpha f(0)}{\alpha!} (x,y)^\alpha
$$

Let's go!

- For $|\alpha| = 0$ we have $\alpha = (0,0)$ and so $\partial^\alpha f(0) = f(0) = 0$, and $\alpha! = 1$.
- For $|\alpha| = 1$ then $\alpha = (1,0)$ or $\alpha = (0,1)$. Call these α_x and α_y respectively, in either case $\alpha_x! = \alpha_y! = 1$ and then:

$$
\begin{aligned}
\partial^{\alpha_x} f(0) &= \frac{\partial f}{\partial x}(0) = 2x \cdot \cos(x^2 + y)\Big|_0 = 0 \\
\partial^{\alpha_y} f(0) &= \frac{\partial f}{\partial y}(0) = \cos(x^2 + yy^2)\Big|_0 = 1
\end{aligned}
$$

- For $|\alpha = 2|$ we have $\alpha_{xx} = (2,0)$ where $\alpha_{x,x}! = 2$ and $\alpha_{xy} = (1,1)$ and $\alpha_{xy}! = 1$.

And then $\alpha_{yy} = (0,2)$ where $\alpha_{yy}! = 2$. Now:

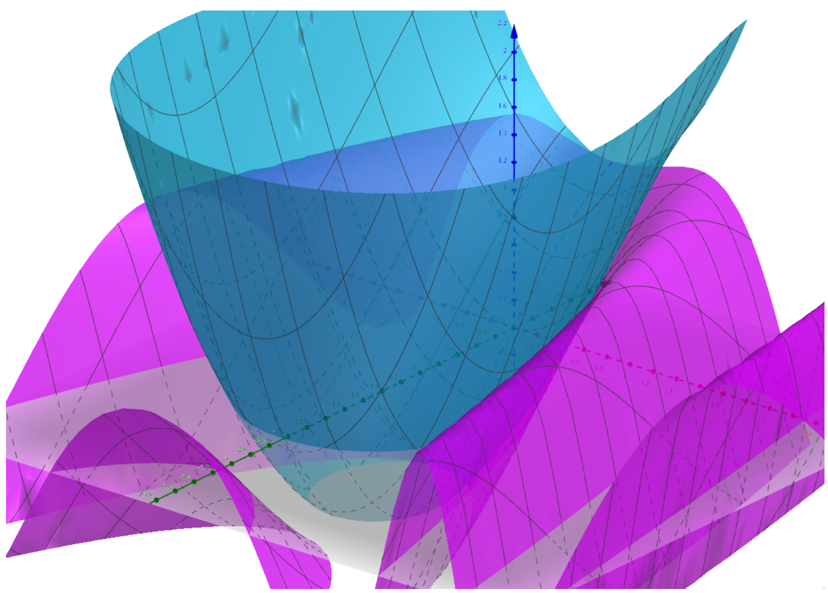
$$\begin{aligned}
\partial^{\alpha_{xx}} f(0) &= \frac{\partial^2 f}{\partial x^2}(0) = 2\cos(x^2+y) - 4x^2 \sin(x^2+y)\Big|_0 = 2 \\
\partial^{\alpha_{xy}} &= \frac{\partial^2 f}{\partial x \partial y}(0) = -2x\sin(x^2+y)\Big|_0 = 0 \\
\partial^{\alpha_{yy}} f(0) &= \frac{\partial^2 f}{\partial y^2}(0) = -\sin(x^2+y)\Big|_0 = 0
\end{aligned}$$

- We omit the case where $|\alpha| = 3$ because we cannot deal. WTF

So then:

$$\begin{aligned}
P(x,y) &= \frac{\partial^{(0,1)} f(0)}{(0,1)!}(x,y)^{(0,1)} + \frac{\partial^{(2,0)} f(0)}{(2,0)!}(x,y)^{(2,0)} + \frac{\partial^{(0,3)} f(0)}{(0,3)!}(x,y)^{(0,3)} \\
&= y + \frac{2}{2}x^2 - \frac{1}{6}y^3 = x^2 + y - \frac{1}{6}y^3
\end{aligned}$$

In the following picture. The blue is our polynomial and the purple is f:



Cool!

Inverse Function Theorem

The inverse function theorem gives a necessary and sufficient condition for a function $f : \mathbb{R}^n \to \mathbb{R}^n$ to be locally invertible with a C^1 inverse.

Definition. *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ where A is open, and let $x_0 \in A$. We say that f is locally invertible around x_0 provided that there is some open neighborhood U of x_0 so that $f\big|_U : U \to f(U)$ is one-to-one, and $f(U)$ is open in $\mathbb{R}^m$ This defines an inverse function $g : f(U) \to U$.*

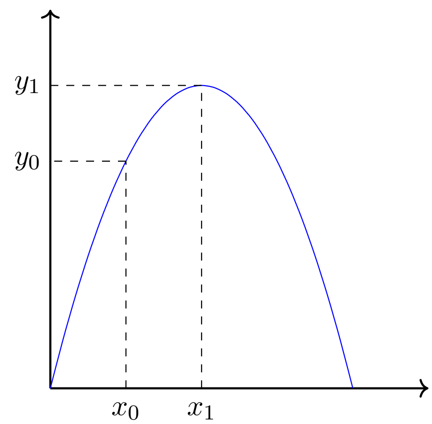
- *We say that f is a local homeomorphism around x_0 provided that both f and g are continuous.*
- *We say that f is a local diffeomorphism around x_0 provided that both f and g are differentiable.*
- *We say that f is a local C^r-diffeomorphism for $r \geq 1$ provided that both f and g are C^r-functions.*
- *We say that f is a locally invertible (resp. homeomorphism, diffeomorphism, C^r-diffeomorphism) provided that it is locally invertible (resp.) around every $x_0 \in A$.*

Remark. Soon we will give an example that is a local diffeomorphism on an open set A but is not a diffeomorphism of A.

Our goal is to find a condition for a function to be a local diffeomorphism. This is easy in one dimension.

The Key Idea

Key Figure



Being a local diffeomorphism neat x is equivalent to being able to express x as a function of y. This means that the graph of $y = f(x)$ can also be regarded as a function $x = g(y)$. This can be done when $\frac{\mathrm{d}f}{\mathrm{d}x}(x_0) \neq 0$. If $\frac{\mathrm{d}f}{\mathrm{d}x}(x_1) = 0$, we might get multiple intersections of lines parallel to the x-axis neat $y = f(x_1)$, which means that the graph cannot define a function $x = g(y)$. The inverse function theorem will generalize this intuition to higher dimensions.

Necessity that $Df(x_0)$ is invertible

Proposition. *Suppose that $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ where A is open. Let $x_0 \in A$ and suppose f is differentiable in A. Assume that f is a local diffeomorphism around x_0 and suppose $g : \mathscr{O} \subseteq \mathbb{R}^n \to B(x_0, \delta)$ where $\mathscr{O}$ is open containing $y = f(x_0)$ is the inverse function. Then $Df(x_0)$ is invertible and:*

$$Dg(y_0) = [Df(x_0)]^{-1}$$

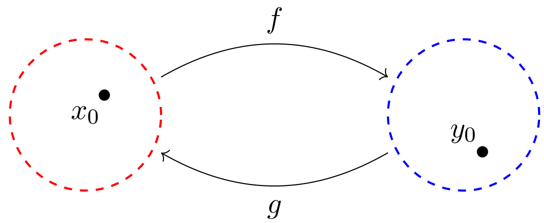
Proof. Consider that:

$$g \circ f : B(x_0, \delta) \to B(x_0, \delta)$$

And $(g \circ f)(x) = x$. Deriving both sides and using the chain rule:

$$Dg(f(x_0))Df(x_0) = Dg(y_0)Df(x_0) = \mathrm{Id}$$

And so $Df(x_0)$ is invertible and $Dg(y_0) = [Df(x_0)]^{-1}$.



Remark. The above proposition shows us that we cannot have a local diffeomorphism as defined from $A \subseteq \mathbb{R}^n \to \mathbb{R}^m$.

One can ask if this is also the case for local homeomorphism. The answer is yes. However, the proof is more involved and uses tools from algebraic topology (Brouwer's invariance of domain theorem)

Handout 6

Second Try: Jordan measure

- **Definition of Jordan measure.** The main caveat of elementary measure is that it only allows us to measure elementary sets, which is a fairly restrictive family of sets. Building on the old intuition (going back at least to Archimedes) we can lower bound (respectively upper bound) the measure of a set by approximating it from within (respectively without) by an elementary set, i.e. if A and B are elementary and $A \subset E \subset B$, then the measure of E (if it exists) should be sandwiched between that of A and B.

  Definition 0.1 (Jordan measure)**.** Let $E \subset \mathbb{R}^d$ be a bounded set.

  – The *Jordan inner measure* $\underline{m}_J(E)$ of E is defined as
  $$\underline{m}_J(E) = \sup_{A \subset E, A \text{ elementary}} m(A).$$
  Here $m(A)$ is the elementary measure of A.

  – The *Jordan outer measure* $\overline{m}_J(E)$ of E is defined as
  $$\overline{m}_J(E) = \inf_{A \supset E, A \text{ elementary}} m(A).$$

  – If $\underline{m}_J(E) = \overline{m}_J(E)$, we say that E is Jordan measurable, and call the common value $m(E)$ (the Jordan measure of E).
  By convention, we do not consider unbounded sets to be Jordan measurable.

Q1) Assume that $E \subset \mathbb{R}^d$ is bounded. Show that the following are equivalent:

a) E is Jordan measurable.

b) For every $\epsilon > 0$, there exists elementary sets A and B such that $A \subset E \subset B$ and $m(B \setminus A) \leq \epsilon$.

c) For every $\epsilon > 0$, there exists an elementary set A such that $\overline{m}_J(E \Delta A) \leq \epsilon$.

Q2) Deduce that every elementary set E is Jordan measurable and that its Jordan measure is the same as its elementary measure. In particular, $m(\emptyset) = 0$.

- **Properties of Jordan measure** Let E, F be Jordan measurable sets.

Q3) Clearly $m(E) \geq 0$. Show that

(a) Show that $E \cup F, E \cap F, E \setminus F$, and $E \Delta F$ are all Jordan measurable.

(b) (Finite additivity) If E and F are disjoint, then $m(E \cup F) = m(E) + m(F)$.

(c) (Monotonicity) If $E \subset F$, then $m(E) \leq m(F)$.

(d) (Finite subadditivity) $m(E \cup F) \leq m(E) + m(F)$.

(e) (Translation invariance) for any $x \in \mathbb{R}^d$, $m(E + x) = m(E)$.

- **Some Jordan measurable sets.** Let B be a closed box of $\mathbb{R}^d$ and $f : B \to \mathbb{R}$ a continuous function.

Q5) Show that the graph $\{(x, f(x)) : x \in B\} \subset \mathbb{R}^{d+1}$ is Jordan measurable in $\mathbb{R}^{d+1}$ and that it has Jordan measure 0. *Hint: Use that f is uniformly continuous.*

Q6) Show that the set $\{(x, t) : x \in B, 0 \leq t \leq f(x)\} \subset \mathbb{R}^{d+1}$ is Jordan measurable.

From this we conclude that some familiar sets like triangles in $\mathbb{R}^2$ and balls in $\mathbb{R}^d$ are Jordan measurable.

MATH 395 Notes

Faye Jackson

October 9, 2020

Exercise 1.

Proof. Let's go!

(a $\implies$ b) Fix a Jordan mesaurable set E and some $\varepsilon > 0$. By definition of suprema and infima there exist elementary sets $A \subseteq E$ and $E \subseteq B$ such that:

$$m(E) - \frac{\varepsilon}{2} < m(A) \leq m(E) \leq m(B) < m(E) + \frac{\varepsilon}{2}$$

Note now that $A \cup (B \setminus A) = B$ since $A \subseteq B$ and so since this union is disjoint:

$$m(B) = m(A) + m(B \setminus A) < m(E) + \frac{\varepsilon}{2}$$

So then:

$$m(B \setminus A) < m(E) - m(A) + \frac{\varepsilon}{2} < \varepsilon$$

Great! Thus c) holds.

(b $\implies$ c) Fix a set E satisfying the condition in (b). Now fix $\varepsilon > 0$. There must be elementary sets $A \subseteq E \subseteq B$ so that $m(B \setminus A) \leq \varepsilon$. Note that:

$$E \triangle A = (E \setminus A) \cup (A \setminus E) = E \setminus A$$

So then note that $E \setminus A \subseteq B \setminus A$ since $E \subseteq B$. Also $B \setminus A$ is an elementary set, so we must have by definition of infimum that:

$$\overline{m}_J(E \triangle A) \leq m(B \setminus A) \leq \varepsilon$$

And so we are done!

(b $\implies$ a) Fix some set E satisfying (b). In order to show that $\overline{m}_J(E) = \underline{m}_J(E)$ we will show that for all $\varepsilon > 0$ we have $|\overline{m}_J(E) - \underline{m}_J(E)| \leq \varepsilon$. Fix some $\varepsilon > 0$, then there exists elementary sets $A \subseteq E \subseteq B$ so that $m(B \setminus A) \leq \varepsilon$. Note that we must have by previous work and definitions that:

$$\begin{aligned} m(A) &\leq m(B) \\ m(B) &= m(A) + m(B \setminus A) \\ \overline{m}_J(E) &\leq m(B) \\ \underline{m}_J(E) &\geq m(A) \\ \overline{m}_J(E) - \underline{m}_J(E) &\leq m(B) - m(A) \\ &= m(B \setminus A) \leq \varepsilon \end{aligned}$$

Now note that for every elementary sets C_1 and C_2 with $C_1 \subseteq E \subseteq C_2$ we must have $m(C_1) \leq m(C_2)$. This shows by 295 that:

$$\overline{m}_J(E) = \sup_{\substack{C \subseteq E \\ C \text{ elementary}}} m(C) \leq \inf_{\substack{C \supseteq E \\ C \text{ elementary}}} m(C) \leq \underline{m}_J(E)$$

Therefore we have that:

$$|\overline{m}_J(E) - \underline{m}_J(E)| = \overline{m}_J(E) - \underline{m}_J(E) \leq \varepsilon$$

Taking $\varepsilon \to 0$ we know that the outer Jordan measure agrees with the inner Jordan measure and so E is Jordan measurable.

(c $\implies$ b) Fix some set E satisfying (c). Now fix some $\varepsilon > 0$. There exists some elementary set A with $\overline{m}_J(E \triangle A) \leq \frac{\varepsilon}{2} < \varepsilon$. Therefore by definition of infima there must be some elementary set B so that $E \triangle A \subseteq B$ and:

$$\overline{m}_J(E \triangle A) \leq m(B) < \varepsilon$$

Now note that $E \setminus A \subseteq B$, and so $E \subseteq A \cup B$. Set $D := A \cup B$. Now consider $C := A \setminus B$ and note that:

$$A \setminus B \subseteq A \setminus (E \triangle A) = A \setminus ((A \cup E) \setminus (A \cap E)) = A \setminus (A \cup E) \cup (A \cap E) = A \cap E$$

And therefore $C \subseteq E$. We then note that:

$$D \setminus C = (A \cup B) \setminus (A \setminus B) = B$$

So we know in particular that $m(D \setminus C) = m(B) < \varepsilon$. Since D and C are elementary sets we must have that E satisfies (b).

Exercise 2.

Proof. Fix some elementary set E. We show that E satisfies (b) from Exercise 1 and so E is Jordan measurable. Fix some $\varepsilon > 0$ and note that $E \subseteq E \subseteq E$ and furthermore:

$$m(E \setminus E) = m(\emptyset) = 0 < \varepsilon$$

So we know that E is jordan measurable. We now only to show that the Jordan measure of E agrees with the elementary measure of E. To do this we calulate $\underline{m}_J(E)$. Fix some $A \subseteq E$ with A elementary, by previous homework $m(A) \leq m(E)$ so the elementary measure of E is an upper bound on the set defining $\underline{m}_J(E)$. Furthermore, this upper bound belongs to the set defining $\underline{m}_J(E)$ since $E \subseteq E$ and E is elementary. Therefore it is a maximum for that set, and is thus the supremum.

This gives us that the Jordan measure of E, which is equal to the Jordan inner measure is also equal to the elementary measure of E just as desired. Perfect!

Exercise 1c.

Proof. Let's go! Fix E and F as Jordan measurable sets.

- **TODO**
- **TODO**
- **TODO**
- **TODO**

This gives us exactly what we want.

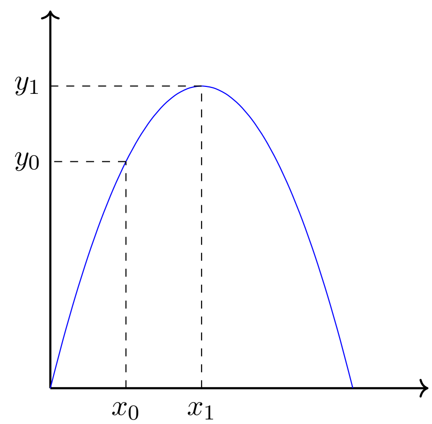
MATH 395 Notes

Faye Jackson

October 12, 2020

Recall. Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ with A open. Let $x_0 \in A$. We say that f is locally invertible near $x_0 \in A$ provided that there eixsts $U, V \subseteq \mathbb{R}^n$ such that $x_0 \in U$, $f(x_0) \in V$, and f is bijective from U to V. Similarly we define local homeomorphism and local diffeomorphisms.

Main Question: When is a function f a local diffeomorphism? If $y = f(x)$ this means, when can we express x as a function of y.



Then clearly we can only express x as a function of y in a neighborhood of y_0 and not y_1. The reason for this difference is $\frac{\mathrm{d}f}{\mathrm{d}x}(x_0) \neq 0$ whereas $\frac{\mathrm{d}f}{\mathrm{d}x}(x_1) = 0$.

This geometric intuition turns out to be true in any dimension if we require $Df(x_0)$ to be invertible instead of just non-zero. Of course this is equivalent to the determinant of $Df(x_0)$ being nonzero.

Recall. Last time, we showed that if f is a local diffeomorphism near x_0 and $g : U \to V$ is the inverse function with $x_0 \in U$ and $y = f(x_0) \in V$, then:

$$Dg(y_0) = [Df(x_0)]^{-1}$$

This shows the necessity of the condition $Df(x_0)$ being invertible for f to be a local diffeomorphism near x_0. The inverse function theorem (IFT) tells us that this is sufficient

Theorem (Inverse Function Theorem, IFT)**.** *Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \to \mathbb{R}^n$ be of class C^r with $r \geq 1$. Suppose that $x_0 \in A$ and $Df(x_0)$ is invertible, then:*

(1) There exists an open neighborhood U of x_0 and an open neighborhood of V of $y_0 = f(x_0)$ such that f is a bijection from U to V

(2) The inverse function $g : V \to U$ is of class C^r as well, and $Dg(y) = [Df(x)]^{-1}$ when $y = f(x)$ for any $x \in U$.

Remark. Another interpretation of IFT is that it allows us to solve an equation:

$$y = f(x)$$

For x in terms of y locally around x_0 when $Df(x_0)$ is invertible. Note that if the funciton f is invertible then $f(x) = Ax$ for some $n \times n$ matrix A, then the ability to solve this equation is exactly the invertibility of A, but $A = Df(x)$ for any x. Wow! The IFT generalizes this to nonlinear functions using differentiability and we work locally.

Remark. The IFT does not guarantee the existence of a global inverse function of $f : A \to \mathbb{R}^n$, but only a local inverse, even if $Df(x)$ is invertible and continuous for all $x \in A$.

The only exception is when $n = 1$, and A is connected. In that case if $f'(x) \neq 0$ and f' is continuous then $f'(x)$ has a definite sign, and so f is either strictly increasing or decreasing. This stops being true for $n \geq 2$

Example. Here's a concrete example. Let $f : A = (1, 2) \times (-\pi, 3\pi) \to \mathbb{R}^2$ where $f(r, \theta) = (r\cos(\theta), r\sin(\theta))$. Then:

$$Df(r, \theta) = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}$$

THen note that:

$$\det(Df(r, \theta)) = r \in (1, 2)$$

And so $Df(r,\theta)$ is invertible on A. However $f(r,0) = (r,0) = f(r,2\pi)$. Thus f is not globally injective, even though the IFT tells us that it is locally

Lemmas for the IFT

Lemma 1. *Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \to \mathbb{R}^n$ be of class C^1. If $Df(x_0)$ is non-singular (that is invertible), then there exists an $\alpha > 0$ and a neighborhood U of x_0 such that:*

$$|f(x) - f(y)| \geq \alpha \, |x - y|$$

For any $x, y \in U$. In particular $f(x) \neq f(y)$ if $x \neq y$. Therefore f is one-to-one on U.

Proof. Let's Go! First we need the linear case:

Let $E = Df(x_0)$. If f were a linear function, that is $f(x) = Ex$, then $f(x) - f(y) = E(x-y)$. Therefore $x - y = E^{-1}(f(x) - f(y))$. This implies that:

$$|x - y| = \left|E^{-1}(f(x) - f(y))\right| \leq \|E^{-1}\| \cdot |f(x) - f(y)|$$

Where we have defined for any matrix $C : \mathbb{R}^n \to \mathbb{R}^m$ the operator norm:

$$\|C\| = \sup_{\substack{x \in \mathbb{R}^n \\ |x| = 1}} |Cx|$$

Great!

Exercise. *Prove that $|Cx| \leq \|C\| \cdot |x|$ for any $x \in \mathbb{R}^n$ and that:*

$$\|C\| \leq nm \cdot \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} |C_{ij}|$$

This is useful for us!

Continuing we then have that:

$$|f(x) - f(y)| \geq \frac{1}{\|E^{-1}\|} \, |x - y|$$

Step 2, we need to generalize. Let $H(x) = f(x) - Ex$ where $E = Df(x_0)$. Then:

$$DH(x) = Df(x) - E$$
$$DH(x_0) = Df(x_0) - E = 0$$

Since H is a C^1 function we can choose $\varepsilon > 0$ so that:

$$\|DH(x)\| \leq \frac{1}{2\|E^{-1}\|}$$

If $x \in B(x_0, \varepsilon)$. Now by the mean value theorem (that is Taylor's Theorem at order 0) we have some c between x and $y$with $x, y \in B(x_0, \varepsilon)$ so that:

$$|H(x) - H(y)| = |DH(c) \cdot (x - y)| \leq \|DH(c)\| \cdot |x - y| \leq \frac{1}{2\|E^{-1}\|} \cdot |x - y|$$

On the other hand:

$$|H(x) - H(y)| = |f(x) - f(y) - E(x - y)| \geq |E(x - y)| - |f(x) - f(y)|$$

Therefore:

$$|f(x) - f(y)| \geq |E(x - y)| - \frac{1}{2\|E^{-1}\|} |x - y|$$

But then by Step 1:

$$\begin{aligned} |f(x) - f(y)| &\geq |E(x - y)| - \frac{1}{2\|E^{-1}\|} |x - y| \\ &\geq \frac{1}{\|E^{-1}\|} |x - y| - \frac{1}{2\|E^{-1}\|} |x - y| = \frac{1}{2\|E^{-1}\|} |x - y| \end{aligned}$$

Exercise. *Suppose $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is C^1, show that the function $x \in A \mapsto \|Df(x)\|$ is continuous. More generally we just need to know that the operator norm is continuous, that is $\mathrm{Mat}(m \times n) \to \mathbb{R}_{\geq 0}$ given by $A \mapsto \|A\|$ is continuous.*

MATH 395 Notes

Faye Jackson

October 14, 2020

More Inverse Function Theorem

Theorem (Inverse Function Theorem, IFT)**.** *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a C^r-function for $r \geq 1$ and suppose $Df(x_0)$ is invertible where $x_0 \in A$. Then f is a local C^r-diffeomorphism around x_0. In other words there are open neighborhoods U of x_0 and V of $f(x_0)$ such that:*

1) f is a bijection from U to V

2) The inverse function $g : V \to U$ is C^r and $Dg(y) = [Df(x)]^{-1}$ where $x \in U$ and $y = f(x)$.

Lemma. *If $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is C^1 and $Df(x_0)$ is non-singular. Then f is locally one-to-one around x_0. More strongly there is an open neighborhood U around x_0 such that for some $\alpha > 0$ we have that for all $x, y \in U$:*

$$|f(x) - f(y)| \geq \alpha\, |x - y|$$

Great!

Lemma. *Suppose $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ (where A is open) is differentiable. If f admits a local minimum (or maximum) at $x_0 \in A$, then $Df(x_0) = 0$.*

Proof. Let $u \in \mathbb{R}^n$ be arbitrary and set $\phi(t) = f(x_0 + tu)$ where $t \in (-\delta, \delta)$ for δ small enough so that $x_0 + tu$ is always in A. Since f has an extremum at x_0, then so does ϕ at 0. By the chain rule ϕ is differentiable on $(-\delta, \delta)$. Therefore $\phi'(0) = 0$, but:

$$\begin{aligned} \phi'(t) &= Df(x_0 + tu) \cdot u \\ 0 = \phi'(0) &= Df(x_0) \cdot u \end{aligned}$$

And this is true for any $u \in \mathbb{R}^n$, so $Df(x_0) = 0$.

Proof of the Inverse Function Theorem, IFT. By the first lemma there exists a neighborhood U of x_0 on which f is one-to-one. By shrinking U if necessary we may also assume that $Df(x)$ is non-singular for every $x \in U$. We may do this because $f \in C^1$ and so Df varies continuously, meaning that since $\det Df(x_0) \neq 0$ we can shrink U to get nonzero determinant all across U. Let $V = f(U)$.

Step 1 We must show V is open in $\mathbb{R}^n$. Take $y \in V$, we want to show that there exists an $\varepsilon > 0$ such that $B(y, \varepsilon) \subseteq V$. Write $y = f(x)$ for some $x \in U$. Since U is open there is some $\delta > 0$ so that $\overline{B(x, \delta)} \subseteq U$. Note that the boundary $\partial B(x, \delta) = \{z \in \mathbb{R}^n \mid |z - x| = \delta\}$ is a compact set, and so if we let $\Gamma = f(\partial B(x, \delta))$ we know that this is compact since f is continuous. Note that $y \notin \Gamma$ because f is one-to-one. Thus there is an $\varepsilon > 0$ such that $B(y, 2\varepsilon) \subseteq \Gamma^c$. We claim that $B(y, \varepsilon) \subseteq V$. To show that, let $c \in B(y, \varepsilon)$ and set:

$$\begin{aligned} \phi : \overline{B(x, \delta)} &\to \mathbb{R} \\ z &\overset{\phi}{\mapsto} |f(z) - c|^2 \end{aligned}$$

Now since ϕ is a continuous function on a compact set it achieves its minimum value at some point $z_\star \in \overline{B(x, \delta)}$. We claim that $z_\star \notin \partial B(x, \delta)$, and so $z \in B(x, \delta)$. Why? Well if $z_\star \in \partial B(x, \delta)$ then $f(z_\star) \in \Gamma$ and so:

$$\begin{aligned} \phi(z_\star) = |f(z_\star) - c|^2 &= |f(z_\star) - y + y - c|^2 \\ &\geq (|f(z_\star) - y| - |y - c|)^2 > (2\varepsilon - \varepsilon)^2 = \varepsilon^2 \end{aligned}$$

This is a problem since $\phi(x) = |y - c|^2 < \varepsilon^2$, but this contradicts the fact that ϕ has its minimum at $z_\star$. Therefore $z_\star \in B(x, \delta)$ since $z \in \overline{B(x, \delta)}$ and $z \notin \partial B(x, \delta)$. By Lemma 2 we must have that $D\phi(z_\star) = 0$ so we calculate the derivative.

Claim. *To justify the above we look at the function $g : \mathbb{R}^n \to \mathbb{R}$ defined by $g(x) = |x|^2$*

Consider that:

$$\begin{aligned}
g(x_1, \ldots, x_n) &= \sum_{i=1}^{n} x_i^2 \\
\partial_i g(x_1, \ldots, x_n) &= 2x_i \\
Dg(x_1, \ldots, x_n) &= (2x_1, \ldots, 2x_n) = 2x
\end{aligned}$$

So then setting $F(z) = f(z) - c$ and so:

$$D\phi(z) = Dg(F(z)) \cdot DF(z) = 2F(z) \cdot Df(z) = 2(f(z) - c) \cdot Df(z)$$

This gives that:

$$0 = D\phi(z_\star) = 2(f(z_\star) - c)Df(z_\star)$$

Since $Df(z_\star)$ is invertible, this implies that $f(z_\star) - c = 0$, and so $f(z_\star) = c$. Therefore $c \in f(B(x, \delta)) \subseteq f(U)$. And so $B(y, \varepsilon) \subseteq f(U) = V$ as desired.

Great! The conclusion of Step 1 is that $f : U \to V$ is one-to-one, onto, and U, V are open. Therefore there exists an inverse function $g : V \to U$ such that $f \circ g = \mathrm{Id}_V$and $g \circ f = \mathrm{Id}_U$.

Step 2: We must show g is continuous. We need to show that $g^{-1}(U')$ is open for every open $U' \subseteq U$. This is equivalent to showing that $f(U')$ is open for any open $U' \subseteq U$. But wait! This is exactly what we did in Step 1 by replacing U by U'.

Step 3: We show that g is differentiable. To do this. Let $y \in V$ where $y = f(x)$ for some $x \in U$. Now let $E = Df(x)$, by hypothesis E is invertible. We will show that:

$$\frac{g(y+k) - g(y) - E^{-1}(k)}{|h|} \to 0 \text{ as } k \to 0$$

This result implies that g is differentiable at y and $Dg(y) = [Df(x)]^{-1}$ where $y = f(x)$. We know that if $|k|$ is small enough then $\overline{B(y, |k|)} \subseteq V$ by openness. Thus there exists some h such that $y + k = f(x + h)$ for some $x + h \in U$. And so we know $k = f(x + h) - f(x)$. Now note that $h = g(y + k) - g(y)$ and so $h \to 0$ as $k \to 0$ by continuity of g. By the differentiability of f at x we know that:

$$\begin{aligned}
r(h) &:= f(x+h) - f(x) - Eh \\
&= k - Eh \\
\frac{r(h)}{|h|} &\to 0 \text{ as } |h| \to 0
\end{aligned}$$

Now we know that:

$$\begin{aligned}
E^{-1}r(h) &= E^{-1}k - h = E^{-1}k - g(y+k) + g(y) \\
\frac{-E^{-1}r(h)}{|k|} &= \frac{g(y+k) - g(y) - E^{-1}k}{|k|}
\end{aligned}$$

It then suffices to show that $\lim_{k \to 0} \frac{E^{-1}r(h)}{|k|} = 0$. It suffices to show that $\lim_{k \to 0} \frac{r(h)}{|k|} = 0$, since E^{-1} is linear. Writing then:

$$\frac{r(h)}{|k|} = \frac{r(h)}{|h|} \frac{|h|}{|k|}$$

Since $\frac{r(h)}{|h|} \to 0$ as $|h| \to 0$ and since $|h| \to 0$ as $|k| \to 0$ it suffices to show that $\frac{|h|}{|k|}$ is bounded by some $C > 0$ for nonzero but small enough k. Recall that:

$$\begin{aligned}
h &= E^{-1}k - E^{-1}r(h) \\
|h| &= \left|E^{-1}(k - r(h)\right| \\
&\leq \|E^{-1}\| \cdot |k - r(h)| \\
&\leq \|E^{-1}\| \cdot (|k| + |r(h)|)
\end{aligned}$$

Now since $\frac{r(h)}{|h|} \to 0$ as $|h| \to 0$ if $|h|$ is small enough we get:

$$\frac{|r(h)|}{|h|} \leq \frac{1}{2\|E^{-1}\|}$$

Therefore if $|k|$ is small enough then $|h|$ is small enough so that $|r(h)| \leq \frac{|h|}{2\|E^{-1}\|}$.

And therefore:

$$
\begin{aligned}
|h| &\leq \|E^{-1}\| \left(|k| + \frac{|h|}{2\|E^{-1}\|} \right) \\
&= \|E^{-1}\| \, |k| + \frac{|h|}{2} \\
|h| &\leq 2\|E^{-1}\| \, |k| \\
\frac{|h|}{|k|} &\leq 2\|E^{-1}\|
\end{aligned}
$$

Pulling this all together:

$$
\left| \frac{g(y+k) - g(y) - E^{-1}k}{|k|} \right| = \left| \frac{E^{-1}r(k)}{|h|} \right|
$$

And we know that:

$$
\left| \frac{r(h)}{|k|} \right| = \frac{|r(h)|}{|h|} \cdot \frac{|h|}{|k|} \leq 2\|E^{-1}\| \frac{|r(h)|}{|h|} \to 0 \text{ as } h \to 0
$$

And so since $h \to 0$ as $k \to 0$ and E^{-1} is linear, we are done, g is differentiable.

Step 4: We need to check that $g \in C^r(V)$. We have shown that $Dg(y) = [Df(x)]^{-1}$ where $y = f(x)$. We can write this as:

$$
Dg = [Df]^{-1} \circ g
$$

By Cramer's rule $[Df]^{-1}$ is a rational function (a polynomial over a polynomial) of the partials $\frac{\partial f_i}{\partial x_j}$, and this rational function has nonzero denominator

Recall. Cramer's rule gives you a formula for the inverse of a matrix C, namely:

$$
C^{-1} = \frac{1}{\det C} \cdot [\operatorname{Adj} C]
$$

We have that $\det C$ is a polynomial in entries of C and:

$$
(\operatorname{Adj} C)_{ij} = \det(C_i^j)
$$

Where C_i^j is the same as C except that we replace the i-th column with $\vec{e}_j$. Of course these are all polynomials in terms of the entries of C.

This implies that $[Df]^{-1}$ belongs to C^{r-1} if $f \in C^r$ because Df belongs to C^{r-1}.

Now consider that:

$$Dg = [Df]^{-1} \circ g \tag{$\star$}$$

Now we know that $g \in C^0$ and so since $[Df]^{-1} \in C^0$ we get $Dg \in C^0$. But then $g \in C^1$. Feeding this into ($\star$) again we get that $Dg \in C^1$ if $r \geq 2$, and so $g \in C^2$. We may do this r times to obtain that $g \in C^r$.

MATH 395 Notes

Faye Jackson

October 16, 2020

Theorem (Mean Value Theorem)**.** *For a differentiable function $H : \mathbb{R}^n \to \mathbb{R}$ we have that for any $x, y \in \mathbb{R}^n$ there is some c on the line segment between x and y so that:*

$$H(y) - H(x) = DH(c) \cdot (y - x)$$

Great!

Proof. Set $\phi(t) : [0, 1] \to \mathbb{R}$ as $\phi(t) = H(x + t(y - x))$. By the single-variable mean value theorem there is some $t \in (0, 1)$ so that:

$$\begin{aligned} \phi(1) - \phi(0) &= \phi'(t) \cdot (1 - 0) \\ H(y) - H(x) &= \phi'(t) \end{aligned}$$

Now by the chain rule, if we set $c := x + t(y - x)$, which is on the line segment:

$$\phi'(t) = DH(c) \cdot (y - x)$$

And so we have the statement of the mean value theorem. Of course, this is just Taylor's Theorem at degree $k = 0$.

How to estimate $R_{x_0,k}(x)$

Now for Taylor's Theorem, how do we estimate $R_{k,x_0}(x)$? This will help us to show the Taylor polynomial is a good approximation. Suppose that $f : A \to \mathbb{R}$ is sufficiently differentiable and that we can show for all $x \in A$ that $|\partial^\alpha f(x)| \leq M_{k+1}$

for $|\alpha| = k+1$. So then:

$$
\begin{aligned}
|(x-x_0)^\alpha| &= \left|\prod_{j=1}^{n}(x_j - x_{0,j})^{\alpha_j}\right| \le \left|\prod_{j=1}^{n}|x-x_0|^{\alpha_j}\right| = |x-x_0|^{|\alpha|} \\
|R_{k,x_0}(x)| &= \left|\sum_{|\alpha|=k+1}\frac{\partial^\alpha f(c_x)}{\alpha!}\right| \le \sum_{|\alpha|=k+1}\frac{M_{k+1}}{\alpha!}\,|x-x_0|^{k+1}
\end{aligned}
$$

Worksheet Time

Handout 7

Jordan measure (Continued)

- **Recall.**

  Definition 0.1 (Jordan measure)**.** Let $E \subset \mathbb{R}^d$ be a bounded set.

  – The *Jordan inner measure* $\underline{m}_J(E)$ of E is defined as

  $$\underline{m}_J(E) = \sup_{A \subset E, A \,\text{elementary}} m(A).$$

  Here $m(A)$ is the elementary measure of A.

  – The *Jordan outer measure* $\overline{m}_J(E)$ of E is defined as

  $$\overline{m}_J(E) = \inf_{A \supset E, A \,\text{elementary}} m(A).$$

  – If $\underline{m}_J(E) = \overline{m}_J(E)$, we say that E is Jordan measurable, and call the common value $m(E)$ (the Jordan measure of E).
  By convention, we do not consider unbounded sets to be Jordan measurable.

  Recall from last time that the Jordan measure extends the notion of elementary measure to more general sets. We also saw that the Jordan measure satisfies Boolean closure properties (if E, F are Jordan measurable sets, the so are $E \cup F, E \cap F, E \setminus F,$), as well as finite additivity (If $E_1, \ldots, E_k$ are disjoint are Jordan measurable, then $m(E_1 \cup \ldots \cup E_k) = m(E_1) + \ldots + m(E_k)$, and translation invariance ($m(E) = m(E + x)$ for $x \in \mathbb{R}^d$).

Q1) Show that the graph $\{(x, f(x)) : x \in B\} \subset \mathbb{R}^{d+1}$ is Jordan measurable in $\mathbb{R}^{d+1}$ and that it has Jordan measure 0. *Hint: Use that f is uniformly continuous.*

Q2) Show that the set $\{(x,t) : x \in B, 0 \le t \le f(x)\} \subset \mathbb{R}^{d+1}$ is Jordan measurable.

From this we conclude that some familiar sets like triangles in $\mathbb{R}^2$ and balls in $\mathbb{R}^d$ are Jordan measurable. For instance,

Q3) Show that the open and closed balls $B(x_0, r) = \{y \in \mathbb{R}^d : |x - y| < r\}$ and $\overline{B(x,r)} = \{y \in \mathbb{R}^d : |y - x| \le r\}$ are both Jordan measurable, and that their Jordan measure is $c_d r^d$ for some constant $c_d > 0$ that only depends on the dimension.

Q4) Establish the bound $\left(\frac{2}{\sqrt{d}}\right)^d \le c_d \le 2^d$.

- **Sets that are not Jordan measurable**. This shows that a lot of familiar subsets of $\mathbb{R}^d$ are Jordan measurable, however many subsets of interest aren't: a) all unbounded subsets are not Jordan measurable, and more importantly b) several interesting bounded sets are not too as the following questions show.

Q5) Let $E \subset \mathbb{R}^d$ be bounded. Show that both E and its closure $\overline{E}$ have the same Jordan outer measure.

Q6) Show that E and its interior E° have the same Jordan inner measure.

Q7) Show that E is Jordan measurable if and only if the topological boundary $\partial E = \overline{E} \setminus E^\circ$ has Jordan outer measure 0.

Q8) Show that the bullet-riddled square $[0,1] \setminus \mathbb{Q}^2$, and the set of bullets $[0,1] \cap \mathbb{Q}^2$ both have Jordan inner measure zero and Jordan outer measure one. In particular, both sets are not Jordan measurable.

MATH 395 Notes

Faye Jackson

October 19, 2020

Announcements

- Midterm is Wednesday (class time)
  - Cameras should be on
  - Be ready 5 minutes earlier
  - Exam from 1pm $\to$ 2:20pm
  - From 2:30 $\to$ 2:30 pm upload your answers to gradescope
- No class on Friday October 30th because Hani has to work with the NSF

Concluding Remarks on the Inverse Function Theorem

The IFT says that if $y = f(x)$ for $x \in \mathbb{R}^n$ satisfies $Df(x_0)$ being non-singular, then there exists an inverse function near x_0. In other words, this means that specifying $(y_1, \ldots, y_n)$ completely determines $(x_1, \ldots, x_n)$ at least locally around x_0 for $y_0 = f(x_0)$.

This means that we can use $(y_1, \ldots, y_n)$ as a coordinate system around x_0 instead of $(x_1, \ldots, x_n)$.

Example. Let $f : [0, \infty) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^3$ be given by:

$$f(\rho, \phi, \theta) = \rho(\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)$$

This is the spherical coordinate system, note that:

$$
\begin{aligned}
Df &= \begin{pmatrix} \sin\phi\cos\theta & \rho\cos\phi\cos\theta & -\rho\sin\phi\sin\theta \\ \sin\phi\sin\theta & \rho\cos\phi\sin\theta & \rho\sin\phi\cos\theta \\ \cos\phi & -\rho\sin\phi & 0 \end{pmatrix} \\
\det Df(\rho,\phi,\theta) &= -\rho\sin\phi\sin\theta(-\rho\sin\theta(\sin^2\phi+\cos^2\phi)) \\
&\quad + \rho\sin\phi\cos\theta(-\rho\cos\theta(\sin^2\phi+\cos^2\phi)) \\
&= \rho^2\sin\phi(\sin^2\theta+\cos^2\theta) = \rho^2\sin\phi
\end{aligned}
$$

Now note that $\det Df \neq 0$ whenever $\rho \neq 0$ and $\sin\phi \neq 0$. That is for any $(\rho_0, \phi_0, \theta_0)$ such that $\phi_0 \neq 0$ and $\phi_0 \neq 0$ there is a neighborhood U of $(\rho_0, \phi_0, \theta_0)$ on which f is a diffeomorphism. In particular, we can use (f_1, f_2, f_3) as coordinates on U.

In this example, the inverse function can be computed using:

$$
\rho = \sqrt{f_1^2 + f_2^2 + f_3^2} \qquad \phi = \arccos\left(\frac{f_3}{\rho}\right) \qquad \theta = \arctan\left(\frac{f_2}{f_1}\right)
$$

Around a point for which $\rho \neq 0$ and $\sin\phi \neq 0$, this holds whenever $f_1^2 + f_2^2 \neq 0$.

The Implicit Function Theorem

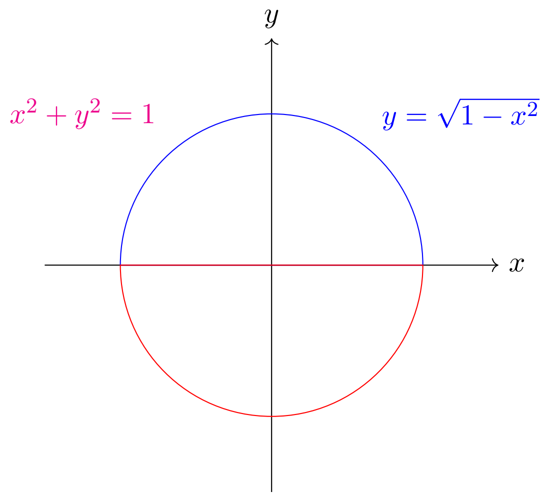
Geometric Motivations

Definition. *A level set of a function $f : \mathbb{R}^n \to \mathbb{R}$ is of the form $\{x \in \mathbb{R}^n \mid f(x) = C\}$ for some constant $C \in \mathbb{R}$*

Consider the function:

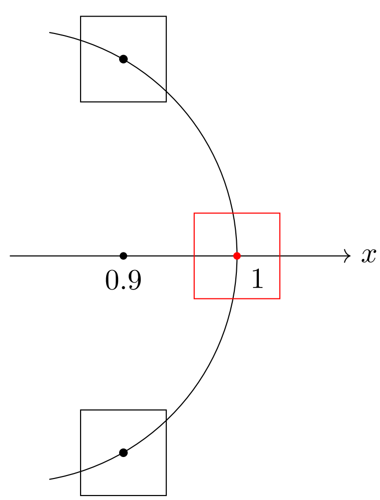
$$
\begin{aligned}
&f : \mathbb{R}^2 \to \mathbb{R} \\
&(x, y) \overset{f}{\mapsto} x^2 + y^2 - 1
\end{aligned}
$$

We know that the equation $f(x, y) = 0$, a level set of f, is the unit circle.



But the upper part of the unit circle is also defined by the function $y = \sqrt{1-x^2}$. In other words, when does the equation $f(x, y) = 0$ define the graph of a function $y = g(x)$. In this case, we say that $f(x, y) = 0$ defines y implicitly in terms of x.

For (a, b) on the unit circle, we can write the equation $f(x, y) = 0$ as $y = g(x)$ in a small neighborhood of (a, b) so long as $(a, b) \neq (1, 0)$ and $(a, b) \neq (-1, 0)$ by the vertical line test



Clearly any red box will violate the vertical line test, and so we can't do this trick near $(1, 0)$.

These are exactly the points where $\frac{\partial f}{\partial y} = 0$. In the context of the implicit function theorem, we are given a function $f(x, y)$ with $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^n$ and

$f : A \subseteq \mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}^n$. When can the level set $\{f = C\}$ locally be described as the graph of a function $y = g(x)$.

Calculus Motivation (Implicit Differentiation)

Suppose that the equation $f(x, y) = 0$ defines y as a function of x (the main assumption). What is $\frac{\mathrm{d}y}{\mathrm{d}x}$. Well:

$$\begin{aligned} f(x, y(x)) &= 0 \\ \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} &= 0 \\ \frac{\mathrm{d}y}{\mathrm{d}x} &= -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} \end{aligned}$$

Indeed the condition $\frac{\partial f}{\partial y} \neq 0$ is again needed to compute $\frac{\mathrm{d}y}{\mathrm{d}x}$. We will see that the Implicit Function Theorem Tells us this is a sufficient condition to be able to express y as a function of x.

Dimension Counting

We would like to find and prove the right generalization of those conditions so that the equation $f(x, y) = 0$ with $x \in \mathbb{R}^k, y \in \mathbb{R}^n$ and $f(x, y) \in \mathbb{R}^p$ can be solved uniquely in terms of x in a sufficiently small neighborhood of (a, b) on the level set.

Let us study the linear problem, i.e. when $f(x, y) = L(x, y)$ and L is a linaer function from $\mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}^p$, that is L is a $p \times (n + k)$ matrix. Write L as:

$$L = \left(\begin{array}{c|c} A & B \end{array} \right)$$

Where A is $p \times k$ and B is $p \times n$. Then $L(x, y) = Ax + By$, and so $L(x, y) = 0$ if and only if $Ax + By = 0$. Therefore y is uniquely solvable in terms of x when $By = Ax$ is uniquely solvable, which happens exactly when B is an invertible matrix. Therefore we must have that $p = n$.

Notice that the matrix B has its columns as $\frac{\partial L(x,y)}{\partial y_j}$ for $1 \leq j \leq n$. This motivates the following:

Definition. *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be differentiable and let $f_1, \ldots, f_m$ be the components of f. We denote:*

- *First:*

$$Df = \frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_n)} = \frac{\partial f}{\partial \vec{x}}$$

  The matrix whose columns are $\frac{\partial f}{\partial x_j} \in \mathbb{R}^m$ *for* $1 \leq j \leq n$*.*

- *Now suppose that* $(x_1, \dots, x_n) = (y, z)$ *for* $y \in \mathbb{R}^k$ *and* $y \in \mathbb{R}^{n-k}$*. We denote then:*

$$\frac{\partial f}{\partial \vec{y}} = \frac{\partial(f_1, \dots, f_n)}{\partial(y_1, \dots, y_k)} = \left(\frac{\partial f}{\partial y_j}\right)_{1 \leq j \leq k}$$
$$\frac{\partial f}{\partial \vec{z}} = \frac{\partial(f_1, \dots, f_n)}{\partial(z_1, \dots, z_{n-k})} = \left(\frac{\partial f}{\partial z_j}\right)_{1 \leq j \leq n-k}$$

The Implicit function theorem states (roughly) that given $f : A \subseteq \mathbb{R}^{k+n} \to \mathbb{R}^n$ where $f(v) = f(x, y)$ with $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^n$, then the level set $\{f(x, y) = 0\}$ defines y as a function of x in a neighborhood of any point (a, b) on the level set <u>if</u> $\frac{\partial f}{\partial \vec{y}}$ is non-singular.

Handout 7

Jordan measure (Continued)

- **Recall.**

  Definition 0.1 (Jordan measure)**.** Let $E \subset \mathbb{R}^d$ be a bounded set.

  – The *Jordan inner measure* $\underline{m}_J(E)$ of E is defined as

  $$\underline{m}_J(E) = \sup_{A \subset E, A \text{ elementary}} m(A).$$

  Here $m(A)$ is the elementary measure of A.

  – The *Jordan outer measure* $\overline{m}_J(E)$ of E is defined as

  $$\overline{m}_J(E) = \inf_{A \supset E, A \text{ elementary}} m(A).$$

  – If $\underline{m}_J(E) = \overline{m}_J(E)$, we say that E is Jordan measurable, and call the common value $m(E)$ (the Jordan measure of E).

  By convention, we do not consider unbounded sets to be Jordan measurable.

Recall from last time that the Jordan measure extends the notion of elementary measure to more general sets. We also saw that the Jordan measure satisfies Boolean closure properties (if E, F are Jordan measurable sets, the so are $E \cup F, E \cap F, E \setminus F,$), as well as finite additivity (If $E_1, \ldots, E_k$ are disjoint are Jordan measurable, then $m(E_1 \cup \ldots \cup E_k) = m(E_1) + \ldots + m(E_k)$, and translation invariance ($m(E) = m(E + x)$ for $x \in \mathbb{R}^d$).

Q1) Show that the graph $\{(x, f(x)) : x \in B\} \subset \mathbb{R}^{d+1}$ is Jordan measurable in $\mathbb{R}^{d+1}$ and that it has Jordan measure 0. *Hint: Use that f is uniformly continuous.*

Q2) Show that the set $\{(x,t) : x \in B, 0 \le t \le f(x)\} \subset \mathbb{R}^{d+1}$ is Jordan measurable.

From this we conclude that some familiar sets like triangles in $\mathbb{R}^2$ and balls in $\mathbb{R}^d$ are Jordan measurable. For instance,

Q3) Show that the open and closed balls $B(x_0, r) = \{y \in \mathbb{R}^d : |x-y| < r\}$ and $\overline{B(x,r)} = \{y \in \mathbb{R}^d : |y-x| \le r\}$ are both Jordan measurable, and that their Jordan measure is $c_d r^d$ for some constant $c_d > 0$ that only depends on the dimension.

Q4) Establish the bound $\left(\frac{2}{\sqrt{d}}\right)^d \le c_d \le 2^d$.

- **Sets that are not Jordan measurable**. This shows that a lot of familiar subsets of $\mathbb{R}^d$ are Jordan measurable, however many subsets of interest aren't: a) all unbounded subsets are not Jordan measurable, and more importantly b) several interesting bounded sets are not too as the following questions show.

Q5) Let $E \subset \mathbb{R}^d$ be bounded. Show that both E and its closure $\overline{E}$ have the same Jordan outer measure.

Q6) Show that E and its interior E° have the same Jordan inner measure.

Q7) Show that E is Jordan measurable if and only if the topological boundary $\partial E = \overline{E} \setminus E^\circ$ has Jordan outer measure 0.

Q8) Show that the bullet-riddled square $[0,1] \setminus \mathbb{Q}^2$, and the set of bullets $[0,1] \cap \mathbb{Q}^2$ both have Jordan inner measure zero and Jordan outer measure one. In particular, both sets are not Jordan measurable.

Handout 7

Jordan measure and Riemann Integration

It turns out that the notion of Jordon measurability of sets is intimately related (in a way essentially equivalent) to the notion of Riemann integrability of functions. We will only display this relation in dimension 1.

- **Recall.** To define the Riemann[1] integral of a bounded function f on an interval $[a, b] \subset \mathbb{R}$, we first recall the notion of a partition $\mathcal{P}$ which is a set of points $x_0 = a < x_1 < x_2 < \ldots < x_n = b$, the norm of the partition is $\Delta\mathcal{P} = \max_{1 \leq k \leq n} x_k - x_{k-1}$, and we denote by $\Delta x_k = x_k - x_{k-1}$. For each such partition, we define to quantities:

$$L(f, \mathcal{P}) = \sum_{k=1}^{n} f(x_*)\Delta x_k, \quad and \quad U(f, \mathcal{P}) = \sum_{k=1}^{n} f(x^*)\Delta x_k,$$

where $x_* = \inf_{[x_{k-1}, x_k]} f$ and $x^* = \sup_{[x_{k-1}, x_k]} f$.

Afterwards, we define the lower and upper Darboux integrals respectively as

$$\underline{\int_a^b} f(x)dx = \sup_{\mathcal{P}} L(f, \mathcal{P}), \quad and \quad \overline{\int_a^b} f(x)dx = \inf_{\mathcal{P}} U(f, \mathcal{P}).$$

where the extrema above are taken over all partitions of the interval $[a, b]$. We say that f is Riemann integrable if the above two numbers are equal. We define the common value as the Riemann (or Darboux) integral of f.

[1]Strictly speaking, we are recalling here the notion of Darboux integral, but that is equivalent to the notion of Riemann integrability that is often covered in introductory calculus classes.

Q1) Let $[a, b]$ be an interval and let $f : [a, b] \to \mathbb{R}$ be a bounded nonnegative function. Show that f is Riemann integrable if and only if the set $E := \{(x, t) : x \in [a, b] : 0 \leq t \leq f(x)\}$ is Jordan measurable in $\mathbb{R}^2$.

Q2) Let $[a, b]$ be an interval and let $f : [a, b] \to \mathbb{R}$ be a bounded function. Show that f is Riemann integrable if and only if the sets $E_+ := \{(x, t) : x \in [a, b] : 0 \leq t \leq f(x)\}$ and $E_- := \{(x, t) : x \in [a, b] : f(x) \leq t \leq 0\}$ are Jordan measurable in $\mathbb{R}^2$.

Remark. The above results generalize to higher dimensions. For that we will need a notion of Riemann (or Darboux) integrability on $\mathbb{R}^d$ $(d \geq 2)$. We will discuss this theory in our lectures, starting next week.

MATH 395 Notes

Faye Jackson

October 23, 2020

Worksheet 7

Proof Sketch of Q3. We talk about this by doing induction. Clearly any ball of radius r in one dimension is measurable, since this will just be a line.

Fix $d \in \mathbb{N}$ so that $B(0, r) \subseteq \mathbb{R}^d$ is measurable. We will show that $B(0, r) \subseteq \mathbb{R}^{d+1}$ is measurable. We consider the following function defined on the box $[-r, r]^d$:

$$f : [-r, r]^d \to \mathbb{R}^{d+1}$$
$$f(x) = \begin{cases} \sqrt{1 - \|x\|^2} & \text{if } \|x\| < r \\ 0 & \text{otherwise} \end{cases}$$

This will give a hemisphere of $B(0, r) \subseteq \mathbb{R}^{d+1}$, and we can glue two of these together to give the full ball. We then can take away the graphs of the functions and we will win.

Proof Sketch of Q5. Fix some bounded subset $E \subseteq \mathbb{R}^d$. We will show that E and $\overline{E}$ have the same Jordan outer measure. To do this lets show that $\overline{m}_J(E) \leq \overline{m}_J(\overline{E})$ and $\overline{m}_J(E) \geq \overline{m}_J(\overline{E})$. Let's go!

Lemma. *The closure of any elementary set A has the same elementary measure as A, and in fact $\overline{A}$ is an elementary set.*

Proof. First note that clearly $m(A) \leq m(\overline{A})$ by monotonicity. Write A as a disjoint union of a finite number of boxes $B_1, \ldots, B_n$. Now note that:

$$\overline{A} = \overline{\left(\bigcup_{k=1}^{n} B_k \right)} = \bigcup_{k=1}^{n} \overline{B}_k$$

We will justify this second equality:

($\subseteq$) Note that $\overline{A}$ is the smallest closed set that contains A. Now note that $\bigcup_{k=1}^n \overline{B}_k$ is closed and since $B_k \subseteq \overline{B}_k$ it contains A. Therefore $\overline{A} \subseteq \bigcup_{k=1}^n \overline{B}_k$.

($\supseteq$) Fix some $x \in \bigcup_{k=1}^n \overline{B}_k$. Then $x \in \overline{B}_j$ for some $1 \leq j \leq n$. Therefore since $B_j \subseteq \bigcup_{k=1}^n B_k$ that we must have $x \in \overline{B}_j \subseteq \overline{\bigcup_{k=1}^n B_k} = \overline{A}$.

Now note that B_k is a box, and so when we take its closure that is still a box, and all the intervals making up the product become closed intervals. This does not change the measure, and so $m(B_k) = m(\overline{B}_k)$. The union above demonstrates that $\overline{A}$ is elementary and by finite subadditivity:

$$m(\overline{A}) \leq \sum_{k=1}^n m(\overline{B}_k) = \sum_{k=1}^n m(B_k) = m(A)$$

And so we must have since $m(A) \leq m(\overline{A})$ that $m(A) = m(\overline{A})$.

Fix some elementary set A that contains $\overline{E}$, this must exist since E is bounded, and thus $\overline{E}$ is bounded. Then A clearly contains E. And so $\overline{m}_J(E) \leq m(E)$. This shows $\overline{m}_J(E)$ is a lower bound for the set definining $\overline{m}_J(\overline{E})$. By the definition of infimum then $\overline{m}_J(E) \leq \overline{m}_J(\overline{E})$.

Now fix some elementary set A that contains E, this must exist since E is bounded. Then $\overline{A}$ contains $\overline{E}$, and so $\overline{m}_J(\overline{E}) \leq m(\overline{A}) = m(A)$ by the lemma. But then $\overline{m}_J(\overline{E})$ is a lower bound for the set defining $\overline{m}_J(E)$. This means that $\overline{m}_J(\overline{E}) \leq \overline{m}_J(E)$

Therefore $\overline{m}_J(E) = \overline{m}_J(\overline{E})$ and we are done! Great!

Proof Sketch of Q6. This is very similar to Question 5!!! Lets show that E and its interior E° have the same Jordan inner measure! For this we a lemma:

Lemma. *The interior of any elementary set A is elementary and has the same measure as A.*

Proof. First note that $A^\circ \subseteq A$ so by monotonicity if A° is elementary then **TODO**

TODO

Proof Sketch of Q7. Let's go both ways!!!

($\Rightarrow$) Suppose E is Jordan measurable. Then by Q5 and Q6:

$$m(E) = \overline{m}_J(E) = \overline{m}_J(\overline{E})$$
$$m(E) = \underline{m}_J(E) = \underline{m}_J(E^\circ)$$

Now to compute $\overline{m}_J(\partial E)$ we know that $0 \leq \overline{m}_J(\partial E)$ because for any elementary set A we know $0 \leq m(A)$. By the characterization of infima it suffices to find for every $\varepsilon > 0$ some elementary set C containing ∂E so that:

$$0 \leq m(C) \leq \varepsilon$$

Note by characterization of suprema and infima for $\overline{E}$ and E° we have an elementary set A containing $\overline{E}$ and an elementary set B contained in E° so that:

$$m(E) \leq m(A) \leq m(E) + \frac{\varepsilon}{2}$$
$$m(E) - \frac{\varepsilon}{2} \leq m(B) \leq m(E)$$

Now note that $A \setminus B$ contains ∂E since A contains $\overline{E}$ and everything we are cutting from A is in $B \subseteq E^\circ$. Now we know that $A \setminus B$ is elementary, so set $C := A \setminus B$ and we will show $m(C) \leq \varepsilon$. This is simple since:

$$\begin{aligned} B \subseteq E^\circ \subseteq E \subseteq \overline{E} &\subseteq A \\ A &= A \setminus B \sqcup B \\ m(A) &= m(A \setminus B) + m(B) \\ m(C) &= m(A) - m(B) \\ &\leq m(E) + \frac{\varepsilon}{2} - m(E) + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

And so we are done! Great!

($\Leftarrow$) Now suppose that ∂E has outer measure 0. We must show that E is Jordan measurable. To do

By Q5 and Q6 it suffices to show that $\overline{m}_J(\overline{E}) \geq \underline{m}_J(E^\circ)$ and likewise $\overline{m}_J(\overline{E}) \leq \underline{m}_J(E^\circ)$, since these are the outer and inner measures of E respectively.

Fix some elementary set A which contains $\overline{E}$. Then since $\overline{E} \supseteq E \supseteq E^\circ$ we know A contains E°. Now fix an elementary set B so that $B \subseteq E^\circ$. Then $m(B) \leq m(A)$ by monotonicity, so by definition of supremum $\underline{m}_J(E^\circ) \leq m(A)$. Then by definition of infimum $\underline{m}_J(E^\circ) \leq \overline{m}_J(\overline{E})$.

We will prove this one by showing that for every $\varepsilon > 0$ we have:

$$\underline{m}_J(E^\circ) + \varepsilon > \overline{m}_J(\overline{E})$$

And so we get the result by taking ε to 0.

TODO

Worksheet 8

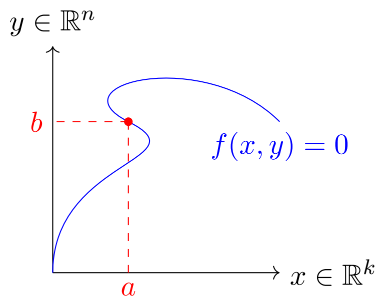
MATH 395 Notes

Faye Jackson

October 26, 2020

More Implicit Function Theorem

Problematic: We have $f : A \subseteq \mathbb{R}^{k+n} \to \mathbb{R}^n$ with $f = f(x, y)$ with $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^n$. We are interested in the level set $L = \{f(x, y) = 0\}$.



Suppose that (a, b) is on the level set, that is $f(a, b) = 0$. Now the equation $f = 0$ gives us n-equations in x

Question: Can we write the condition that $\{f(x, y) = 0\}$ near (a, b) as the graph of a function $y = g(x)$, i.e. $(x, y) \in L$ if and only if $y = g(x)$. In other words, can we solve the system of equations $f(x, y) = 0$ near (a, b) for y in terms of x? In yet other words, does the equation $\{f = 0\}$ define y implicitly in terms of x

Roughly speaking, the implicit function theorem says that the answer is yes provided that $\frac{\partial f}{\partial y} = \frac{\partial(f_1,\ldots,f_n)}{\partial(y_1,\ldots,y_n)}$ is non-singular.

Before stating the theorem precisely, let's state an easier result about the derivative of the implicit function:

Theorem (Implicit Differentiation)**.** *Let $A \subseteq \mathbb{R}^{k+n}$ be open and $f : A \to \mathbb{R}^n$ be differentiable and write $f = f(x, y)$ with $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^n$. Now suppose that the equation $f(x, y) = 0$ defines y implicitly, i.e. there exists a function $g : B \to \mathbb{R}^n$*

defined on an open subset B of $\mathbb{R}^k$ such that $(x, g(x)) \in A$ and $f(x, g(x)) = 0$ for all $x \in B$.

THEN, for $x \in B$ we have:

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

In particular, if $\frac{\partial f}{\partial y}(x, g(x))$ is invertible, then:

$$Dg(x) = -\left[\frac{\partial f}{\partial y}(x, g(x))\right]^{-1} \frac{\partial f}{\partial x}(x, g(x))$$

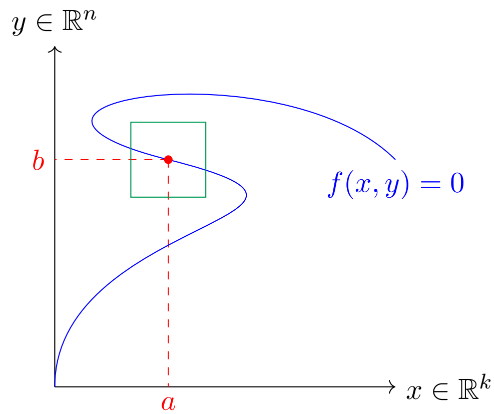
Proof. Then let $h : B \to \mathbb{R}^{k+n}$ be the function $h(x) = (x, g(x))$ then $f \circ h = 0$ by supposition. Take the derivative of this expression, and so by the chain rule:

$$\begin{aligned}
Df(h(x)) \cdot Dh(x) = 0 & \\
Dh(x) &= \left(\frac{I_k}{Dg}\right)\Big|_x \\
Df &= \left(\ \tfrac{\partial f}{\partial x} \ \middle| \ \tfrac{\partial f}{\partial y} \ \right) \\
Df(h(x))Dh(x) &= \left(\ \tfrac{\partial f}{\partial x} \ \middle| \ \tfrac{\partial f}{\partial y} \ \right)\Big|_{h(x)} \cdot \left(\frac{I_k}{Dg}\right)\Big|_x \\
&= \frac{\partial f}{\partial x}(h(x)) + \frac{\partial f}{\partial y}(h(x))Dg(x) = 0
\end{aligned}$$

And this is what we wished to show.

The implicit function theorem tells us that the invertibility of $\frac{\partial f}{\partial y}$ is sufficient for the condition of the above theorem to hold

Theorem (Implicit Function Theorem)**.** *Let $A \subseteq \mathbb{R}^{k+n}$ be open and $f : A \to \mathbb{R}^n$ be of class C^r with $r \geq 1$. Write f in the form $f(x, y)$ with $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^n$. Suppose that $(a, b) \in A$ such that $f(a, b) = 0$.*



If $\frac{\partial f}{\partial y}(a, b)$ is non-singular, then there exists a neighborhood $B \subseteq \mathbb{R}^k$ of a and a unique continuous function $g : B \to \mathbb{R}^n$ such that $g(a) = b$ and $f(x, g(x)) = 0$ for $x \in B$. The function g will in fact be of class C^r. In fact inside the green window, $f(x, y) = 0$ if and only if $y = g(x)$.

Remark. Of course, the variables y for which we solve for in terms of x don't have to be the last n coordinates. They can be any n of the $(n + k)$ coordinates.

Proof. Step 1 (An Auxiliary Function): Consider the auxiliary function:

$$F : A \subseteq \mathbb{R}^{k+n} \to \mathbb{R}^{k+n}$$

$$(x, y) \xmapsto{F} \begin{pmatrix} x \\ f(x, y) \end{pmatrix}$$

$$DF(x, y) = \begin{pmatrix} DF_1 \\ DF_2 \\ \vdots \\ DF_{k+n} \end{pmatrix} = \left(\frac{I_k}{Df} \right)$$

$$= \left(\begin{array}{c|c} I_k & 0 \\ \hline \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{array} \right)$$

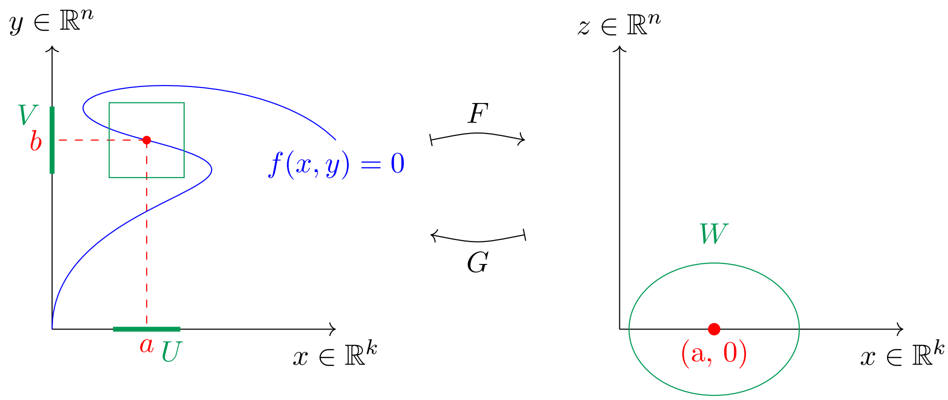
Therefore using block matrices you can check that:

$$\det DF(x,y) = \det I_k \det\left(\frac{\partial f}{\partial y}\right) = \det\left(\frac{\partial f}{\partial y}\right)$$

But we know that $\frac{\partial f}{\partial y}$ is nonsingular at (a,b) and so:

$$\det DF(a,b) = \det \frac{\partial f}{\partial y}(a,b) \neq 0$$

Thus $DF(a,b)$ is nonsingular, and so by the inverse function theorem there exists a neighborhood $U \times V$ of (a,b) such that $a \in U$ is open in $\mathbb{R}^k$ and $b \in V$ is open in $\mathbb{R}^n$ as well as a neighborhood W of $(a,0) \in \mathbb{R}^{k+n}$ such that F is a C^r-diffeomorphism from $U \times V$ onto W.



Let $G : W \to U \times V$ be the the inverse function of F. I.e. $(x,y) = G(x, f(x,y))$. for all $(x,y) \in U \times V$ and $(x,z) = F \circ G(x,z)$ for $(x,z) \in W$. This tells us that G is the identity on its first k coordinate functions. Let $h : W \to V$ be defined as $h(x,z) = (G_{k+1}, G_{k+2}, \ldots, G_{k+n})$, h is clearly C^r since G is C^r by the inverse function theorem.

Step 2 (Definition of g): Let B be a ball around a such that $B \subseteq U$ and $B \times \{0\} \subseteq W$. Now notice that $(x,y) \in B \times V$ satisfies $f(x,y) = 0$ if and only if $F(x,y) = (x,0)$ if and only if $(x,y) = G(x,0) = (x, h(x,0))$. Defining $g(x) = h(x,0)$ for $x \in B$ we have that $(x,y) \in B \times V$ satisfying $f(x,y) = 0$ if and only if $y = g(x)$ for $x \in B$. Clearly g is C^r since h is C^r.

Also note that $(a,b) = G(a,0) = (a, h(a,0))$, and so $b = h(a,0) = g(a)$ as desired.

Step 3 (Uniqueness of g): Suppose that $g' : B \to \mathbb{R}$ is another continuous function that satisfies the conclusions of the theorem. Let $S = \{x \in B \mid g(x) = g'(x)\}$. Clearly since g and g' are continuous, S is closed relative to B. Also, we must have that $a \in S$, since $b = g(a) = g(a')$. We will show that S is also open in B, which would mean that $S = B$, since B is connected. This will finish the proof. We'll leave this until next time

MATH 395 Notes

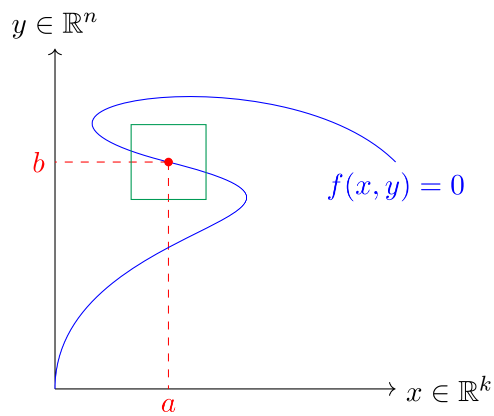
Faye Jackson

October 28, 2020

Note: No class on Friday

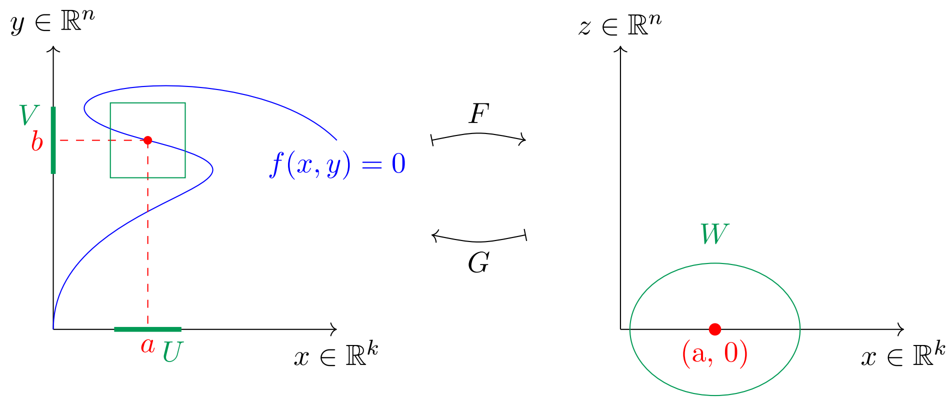
The Proof of the Implicit Function Theorem

Continued Proof of the Implicit Function Theorem. We had an $A \subseteq \mathbb{R}^{k+n}$ and an $f : A \to \mathbb{R}^n$ of class C^r with $r \geq 1$. We also had $f(a,b) = 0$ and $\frac{\partial f}{\partial y}(a,b)$ is nonsingular. We model this with the picture:



We constructed a neighborhood B around a, a neighborhood V around b, and a function $g : B \to V$ satisfying $g(a) = b$ and $f(x,y) = 0$ if and only if $y = g(x)$ for $(x,y) \in B \times V$. We did this with the following steps:

1) We defined an auxiliary function $F(x,y) = (x, f(x,y)) : A \to \mathbb{R}^{k+n}$. We showed that $DF(a,b)$ is invertible since $\frac{\partial f}{\partial y}(a,b)$ is invertible. We then applied the Inverse Function Theorem. This gave us the following picture



We then showed the inverse function $G(x, z)$ must be given as $(x, h(x, z))$ where $h \in C^r$.

2) We then defined g with a neighborhood $B \subseteq U$ such that $B \times \{0\} \subseteq W$. We then defined $g : B \to V$ as $g(x) = h(x, 0)$. This satisfies the desired conditions.

3) We showed the Uniqueness of g. We supposed that $g' : B \to V$ was another continuous function such that $g'(a) = b$ and $f(x, g'(x)) = 0$. We defined $S = \{x \in B \mid g'(x) = g(x)\}$. We want to show that $S = B$. Using the connectedness of B we simply need to show that S is a nonempty subset of B that is both open and closed in B.

   S is clearly nonempty since $g'(a) = g(a)$, and thus $a \in S$. We know S is closed since g, g' are both continuous, and we can rewrite $S = (g - g')^{-1}(\{0\})$. It remains to show that S is open

Let's show this! Let $x_0 \in S$, then $g'(x_0) = g(x_0) \in V$ is open. There must exist a neighborhood B' of x_0 such that $g'(B') \subseteq V$ using the fact that g' is continuous. But then:

$$f(x, g'(x)) = 0 \qquad\qquad x \in B' \subseteq B \qquad\qquad g'(x) \in V$$

But then this must mean that:

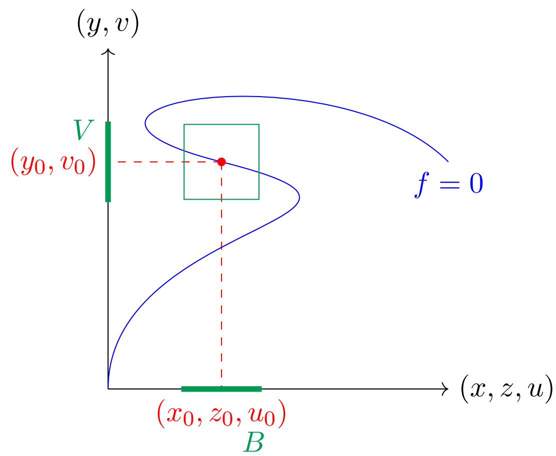
$$\begin{aligned} F(x, g'(x)) &= (x, f(x, g'(x))) = (x, 0) \\ (x, g'(x)) &= G(x, 0) = (x, h(x, 0)) = (x, g(x)) \end{aligned}$$

This of course implies that $g'(x) = g(x)$ for all $x \in B'$. Therefore S is open in B, and we win!!!! Yay ☺

How to Apply the Implicit Function Theorem

Suppose that $f : A \subseteq \mathbb{R}^5 \to \mathbb{R}^2$ is a function in C^r and the equation $f(x, y, z, u, v) = 0$ gives us two equations in five unknowns, and thus by dimension counting, the solution set is a set parameterized in three variables. We expect (under appropriate conditions) that we can solve for two of the variables in terms of the others.

Suppose one wishes to solve for (y, v) in terms of (x, z, u) near a point $(x_0, y_0, z_0, u_0, v_0 = 0$. All we need to check is that $\frac{\partial f}{\partial(y,v)}$ is nonsingular at $(x_0, y_0, z_0, u_0, v_0)$. The implicit function theorem then tells us that we can write $y = \phi(x, z, u)$ and $v = \psi(x, z, u)$



Moreover by implicit differentiation:

$$\frac{\partial(\phi, \psi)}{\partial(x, z, u)}(x_0, z_0, u_0) = -\left[\frac{\partial f}{\partial(y, v)(x_0, y_0, z_0, u_0, v_0)}\right]^{-1} \frac{\partial f}{\partial x}(x, g(x))$$

Example. Show that the system of equations:

$$x^3 - y^3 + z^2 = 0$$
$$z\cos(\pi x) + \sin(\pi y) = 0$$

admits a one-parameter family of solutions around the point $(1, 1, 0)$

Define $f : \mathbb{R}^3 \to \mathbb{R}^2$ by:

$$f(x, y, z) = \begin{pmatrix} x^3 - y^3 + z^2 \\ z\cos \pi x + \sin \pi y \end{pmatrix}$$

Then $f(1,1,0) = 0$ and:

$$Df = \begin{pmatrix} 3x^2 & -3y^2 & 2z \\ -\pi z \sin \pi x & \pi \cos \pi y & \cos \pi x \end{pmatrix}$$
$$Df(1,1,0) = \begin{pmatrix} 3 & - & 0 \\ 0 & -\pi & -1 \end{pmatrix} = \frac{\partial f}{\partial(x,y,z)}$$
$$\frac{\partial f}{\partial(x,z)} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$$

This is of course a non-singular matrix, and so we can solve for (x,z) in terms of y near the point $y = 1$. That is there are functions $\phi, \psi : B \to \mathbb{R}^2$ where B is an open neighborhood of $y = 1$ such that $f(\phi(y), y, \psi(y)) = 0$ for all $y \in B$.

In other words, the solution set near $(1,1,0)$ is a one-parameter family of solutions. We will later find out that this means it is a "manifold of dimension one"

With this we have essentially finished differentiation!

Riemann Integration in Higher Dimensions

Definition of the integral

The purpose os this section is to generalize the notion of the Riemann integral to higher dimesnions

Definition. *We will use some concepts from our Friday sections*

1) *Recall that we defined a box $B \subseteq \mathbb{R}^n$ to be the Cartesian product of n intervals $B = I_1 \times I_2 \times \cdots \times I_n$. Generally $I_1, \ldots, I_n$ can be closed, open, or half open.*

However, in what follows, there will be no loss of generality in considering only closed boxes. Thus to simplify notation, we will assume that all boxes are closed unless stated otherwise

Given $B = [a_1, b_1] \times \cdots \times [a_n, b_n]$ we set:

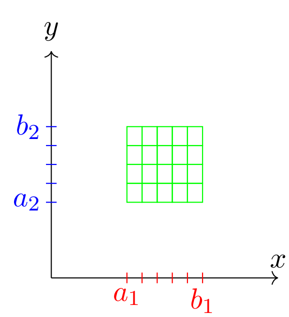
$$m(B) := \prod_{i=1}^{n} (b_i - a_i)$$

2) *Partitions*

$(n = 1)$ *Given an interval $I = [a, b]$ a partition of $[a, b]$ is a finite collection P of points $x_0 = a < x_1 < x_2 < \cdots < x_k = b$. Each $[x_{i-i}, x_i]$ has length $\Delta x_i = x_i - x_{i-1}$. We define the mesh (or norm) of P as:*

$$\|P\| = \max_{1 \le i \le k} \Delta x_i$$

$(n \ge 1)$ *Given a box $B = I_1 \times \cdots \times I_n$, a partition P of B is an n-tuple $(P_1, \ldots, P_n)$ such that P_j is a partition of I_j for each j.*



Each partition P_j decomposes I_j into sub-intervals $I_j^{(1)}, \dots, I_j^{(k_j)}$ with disjoint interiors This gives a decomposition of B into sub-boxes of the form $J_1 \times \cdots J_n$ where $J_j \in \{I_j^{(1)}, \dots, I_j^{(k_j)}\}$.
Notice that the sub-boxes can only intersect at the boundary, that is they have disjoint interiors. The mesh of a partition $P = (P_1, \dots, P_n)$ is $\|P\| = \max_{1\le j\le n} \|P_j\|$.

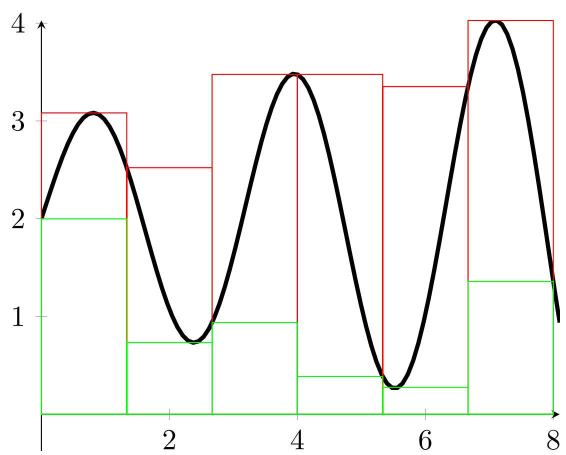
3) *We now define Lower and upper sums. Let B be a box and $f : B \to \mathbb{R}$ be bounded. Let P be a partition of B and denote by $B_1, \dots, B_N$ the resulting subboxes. Let*

$$m_{B_j}(f) := \inf_{x\in B_j} f(x)$$
$$M_{B_i}(f) := \sup_{x\in B_j} f(x)$$

Then we may define the lower and upper sums respectively as:

$$L(f,P) = \sum_{\ell=1}^{N} m_{B_\ell}(f) \cdot v(B_\ell)$$
$$U(f,P) = \sum_{\ell=1}^{N} M_{B_\ell}(f) \cdot v(B_\ell)$$

In one dimension if $f \ge 0$ then $L(f,P)$ is the sum of the green rectangles inscribed by the region under the curve, and $U(f,P)$ is the area of the red rectangles circumscribed by the region under the curve



4) *We define now the Refinement of a partition Let B be a box and le tP =*

$(P_1, \dots, P_n)$ and $Q = (Q_1, \dots, Q_n)$ be two partitions of B. We say that Q is a refinement of P is $P_j \subseteq Q_j$ for every j.

Given two partitions $P = (P_1, \dots, P_n)$ and $P' = (P'_1, \dots, P'_n)$ the common refinment is $Q = (P_1 \cup P'_1, \dots, P_n \cup P'_n)$.

Lemma. *Refining a partition increases lower sums and decreases upper sums. In other words, let P be a partition of a box B and $f : B \to \mathbb{R}$ be bounded. If Q is a refinement of P, then:*

$$L(f, P) \leq L(f, Q) \qquad\qquad U(f, Q) \leq U(f, P)$$

Before proving this lemma, let us state a corollary

Corrolary. *Let B be a box and $f : B \to \mathbb{R}$ be a bounded function. If P and P' are any two partitions of B, then $L(f, P) \leq U(f, P')$.*

Proof of corollary. Clearly for any partition we have $L(f, Q) \leq U(f, Q)$. Le tQ be the common refinement of P and P' and use the lemma to see that:

$$L(f, P) \leq L(f, Q) \leq U(f, Q) \leq U(f, P')$$

Great!

MATH 395 Notes

Faye Jackson

November 2, 2020

Continue Defining the Riemann Integral

Definition. *Given a box $B = [a_1, b_1] \times \cdots \times [a_n, b_n]$ which is closed and a function $f : B \to \mathbb{R}$ that is bounded. We defined a partition $P = (P_1, \ldots, P_n)$ of B as a tuple where each P_j is a partition of $[a_j, b_j]$. We then let $\{B_j\}_{j=1}^N$ be the set of sub-boxes determined by the partition. We then defined the lower sum and upper sum of f over a partition P:*

$$m_{B_j} := \inf_{x \in B_j} f(x)$$

$$M_{B_j} := \sup_{x \in B_j} f(x)$$

$$L(f, P) := \sum_{j=1}^{N} m_{B_j} v(B_j)$$

$$U(f, P) := \sum_{j=1}^{N} M_{B_j} v(B_j)$$

Exercise. $U(f, P) = -L(-f, P)$.

We then talked about refinements of a partition, saying that $Q = (Q_1, \ldots, Q_n)$ is a refinement of $P = (P_1, \ldots, P_n)$ if $P_1 \subseteq Q_1, P_2 \subseteq Q_2, \ldots$.

We defined the common refinement of $P = (P_1, \ldots, P_n)$ and $Q = (Q_1, \ldots, Q_n)$ as $P \cup Q = (P_1 \cup Q_1, \ldots, P_n \cup Q_n)$.

Lemma. *Let P be a partition of a box B and $f : B \to \mathbb{R}$ be bounded. If Q is a* *refinement of P then:*

$$L(f,P) \leq L(f,Q)$$
$$U(f,P) \geq U(f,Q)$$

Proof. We reduce first. Since $U(f,P) = -L(-f,P)$, it is enough to prove the lemma for lower sums.

Now since Q ca be obtained from P by successively adding points to the partition, we can reduce to the case where Q is obtained from $P = (P_1, \ldots, P_n)$ by adding a single point. to P_k for some $1 \leq k \leq n$.

By symmetry, we assume that $k = 1$. Suppose that $B = [a_1, b_1] \times \cdots \times [a_n, b_n]$ and suppose that P_1 consists of the points $a_1 = x_0 < \cdots < x_k = b$. Now Q is obtained by adding the point q that lies in the interior of (x_{p-1}, x_p) for some $1 \leq p \leq k$.

The sub-boxes determined by P are of the form $[x_{i-1}, x_i] \times S$ where S is a subbox of $[a_2, b_2] \times \cdots \times [a_n, b_n]$ determined by the partition $(P_2, \ldots, P_n)$. Let us denote by $\mathscr{S}$ the set of all such subboxes.

The sub-boxes determined by Q are of the form: $[x_{i-1}, x_i] \times S$ for $1 \leq i \leq p-1$ or $p+1 \leq i \leq k$ and $S \in \mathscr{S}$ or $[x_{p-1}, q] \times S$ or $[q, x_p] \times S$ for $S \in \mathscr{S}$. Therefore:

$$\begin{aligned}
L(f,P) &= \sum_{\substack{i=1 \\ S \in \mathscr{S}}}^{k} m_{[x_{i-1},x_i]\times S}(f) \cdot v([x_{i-1}, x_i] \times S) \\
&= \sum_{\substack{i \in \{1,\ldots,p\} \cup \{p+1,\ldots,k\} \\ S \in \mathscr{S}}} m_{[x_{i-1},x_i]\times S}(f) \cdot v([x_{i-1}, x_i] \times S) \\
&\quad + \sum_{S \in \mathscr{S}} m_{[x_{p-1},x_p]\times S}(f) \cdot (x_p - x_{p-1}) \cdot v(S)
\end{aligned}$$

The left sum appears in the definition of $L(f,Q)$, and so we only consider the right sum. The point is that the:

$$\inf_{x \in [x_{p-1},x_p]\times S} f(x) \leq \inf_{x \in [x_{p-1},q]\times S} f(x), \inf_{x \in [q,x_p]\times S} f(x)$$

This implies that:

$$\begin{aligned}
m_{[x_{p-1},x_p]\times S}(f) \cdot (x_p - x_{p-1}) &= m_{[x_{p-1},x_p]\times S}(f) \cdot (q - x_{p-1}) + m_{[x_{p-1},x_p]\times S}(f) \cdot (x_p - q) \\
&\leq m_{[x_{p-1},q]\times S}(f) \cdot (q - x_{p-1}) + m_{[q,x_p]\times S}(f) \cdot (x_p - q)
\end{aligned}$$

But then:

$$
\begin{aligned}
L(f,Q) = &\sum_{\substack{i\in\{1,\dots,p\}\cup\{p+1,\dots,k\} \\ S\in\mathscr{S}}} m_{[x_{i-1},x_i]\times S}(f)\cdot v([x_{i-1},x_i]\times S) \\
&+ \sum_{S\in\mathscr{S}} m_{[x_{p-1},q]\times S}(f)\cdot(q-x_{p-1}) + m_{[q,x_p]\times S}(f)\cdot(x_p-q)
\end{aligned}
$$

And so $L(f,P)\le L(f,Q)$ because:

$$
\begin{aligned}
L(f,P) = &\sum_{\substack{i\in\{1,\dots,p\}\cup\{p+1,\dots,k\} \\ S\in\mathscr{S}}} m_{[x_{i-1},x_i]\times S}(f)\cdot v([x_{i-1},x_i]\times S) \\
&+ \sum_{S\in\mathscr{S}} m_{[x_{p-1},x_p]\times S}(f)\cdot(x_p-x_{p-1})\cdot v(S)
\end{aligned}
$$

And we know that:

$$
\begin{aligned}
&\sum_{S\in\mathscr{S}} m_{[x_{p-1},x_p]\times S}(f)\cdot(x_p-x_{p-1})\cdot v(S) \\
&\le \sum_{S\in\mathscr{S}} m_{[x_{p-1},q]\times S}(f)\cdot(q-x_{p-1}) + m_{[q,x_p]\times S}(f)\cdot(x_p-q)
\end{aligned}
$$

That was disgusting!!!

Corrolary. *If P and P' are any two partitions of B then $L(f,P)\le U(f,P')$. The proof was given last time.*

Definition (Upper integrals, lower integrals, and Riemann integrability)**.** *Let B be a box and let $f : B\to\mathbb{R}$ be a bounded function.*

a) We define the lower and upper integral of f over B respectively as:

$$
\underline{\int_B} f(x)\,\mathrm{d}x = \sup_P L(f,P)
$$
$$
\overline{\int_B} f(x)\,\mathrm{d}x = \inf_P U(f,P)
$$

These numbers exist because $L(f,P)$ is bounded above by $(\sup_{x\in B} f(x))\cdot v(B)$ and $U(f,P)$ is bounded below by $(\inf_{x\in B} f(x))\cdot v(B)$

b) We say that f is Riemann integrable over B provided that the lower and upper integral agree. In this case we define the Riemann integral $\int_B f(x)\,\mathrm{d}x$ as the *common value, aka:*

$$\int_B f(x)\,\mathrm{d}x := \underline{\int_B} f(x)\,\mathrm{d}x = \overline{\int_B} f(x)\,\mathrm{d}x$$

Remark. Strictly speaking, this is the definition of Darboux integrability. The precise definition of Riemann integrability is: A bounded function f is Riemann integrable with integral A on the box B if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that if P is a partition of B with mesh $\leq \delta$, then for any choice of $x_\alpha \in B_\alpha$, where B_α are the sub-boxes determined by P:

$$\left| \sum_{B_\alpha} f(x_\alpha) v(B_\alpha) - A \right| < \varepsilon$$

We will prove these are equivaelnt on Homework 9. F

Remark. Suppose that $f : B \subseteq \mathbb{R}^2 \to \mathbb{R}$ is a non-negative function. Then $L(f, P)$ is the total volume of a bunch of boxes under the graph of f whereas the upper sum is the total volume of a bunch of boxes that are circumscribed

Exercise. *Show that if $f : B \subseteq \mathbb{R}^n \to \mathbb{R}$ is non-negative and bounded. Then f is Riemann integrable if and only if the region in $\mathbb{R}^{n+1}$ under the graph of f given by:*

$$R = \{(x, x_{n+1}) \in \mathbb{R}^n \times \mathbb{R} \mid 0 \leq x_{n+1} \leq f(x)\}$$

is Jordan measurable with $m(R) = \int_B f(x)\,\mathrm{d}x$.

Example. Let $f : [0,1]^2 \to \mathbb{R}$ be defined as:

$$f(x,y) = \begin{cases} 0 & \text{if } x \text{ and } y \text{ are rationally dependent} \\ 1 & \text{otherwise} \end{cases}$$

We call x and y rationally dependent provided that there exists $(k_1, k_2) \in \mathbb{Z}^2$ such that $(k_1, k_2) \neq 0$ and $k_1 x + k_2 y = 0$.

Now let P be any partition of $B = [0,1]^2$. For any subbox R resulting from the partition we have:

$$m_R(f) = \inf_R f = 0$$

$$M_R(f) = \sup_R f = 1$$

Since for any subbox of $[0,1]^2$ with non-empty interior, there exists $(x, y) \in R$ such that both x and y are rational numbers, and so they are rationally dependent. For the second statement, since for any sub-box of $[0,1]^2$ with non-empty interior, there exists $(x, y) \in R$ such that x is a non-zero rational and y is irrational. This implies that x, y are rationally independent.

Therefore:

$$L(f, P) = 0 \qquad\qquad U(f, P) = 1$$

For any partiton P of $[0,1]^2$. And therefore:

$$\underline{\int_B} f(x)\,\mathrm{d}x = 0 \qquad\qquad \overline{\int_B} f(x)\,\mathrm{d}x = 1$$

Therefore, f is not integrable

Theorem 1 (The Riemann Condition)**.** *Let B be a box in $\mathbb{R}^n$ and let $f : B \to \mathbb{R}$ be a bounded function. Then:*

a) We always have that $\underline{\int_B} f(x)\,\mathrm{d}x \leq \overline{\int_B} f(x)\,\mathrm{d}x$

b) f is integrable if and only if for every $\varepsilon > 0$ there exists a partition P of B for which $U(f, P) - L(f, P) < \varepsilon$.

Remark. Reminiscient of the exercise in our discussion sections that E is Jordan measurable if for any $\varepsilon > 0$ there eixsts elementary sets $A \subseteq E \subseteq B$ such that $m(B \setminus A) < \varepsilon$.

Proof. Part (a) is trivial since we saw that $L(f, P) \leq U(f, P')$ for any P and P'. Taking the sup over P and the inf over P' gives the result.

For (b), there are two directions:

($\Rightarrow$) Suppose f is integrable and $\varepsilon > 0$. Choose a partiton P_1 such that:

$$\left| L(f, P_1) - \int_B f \right| < \frac{\varepsilon}{2}$$

and another partition P_2 such that:

$$\left| U(f, P_2) - \int_B f \right| < \frac{\varepsilon}{2}$$

Then we know that $U(f, P_2) - L(f, P_1) < \varepsilon$. Take P to be the common refinement of P_1 and P_2. Then we know that:

$$L(f, P_1) \leq L(f, P) \leq U(f, P) \leq U(f, P_2)$$

This means that $U(f, P) - L(f, P) \leq U(f, P_2) - L(f, P_1) < \varepsilon$. Thus we win!

($\Leftarrow$) Let $\varepsilon > 0$ be arbitrary. Choose a partition P such that $U(f, P) - L(f, P) < \varepsilon$. Then:

$$\overline{\int_B} f - \underline{\int_B} f \leq U(f, P) - L(f, P) < \varepsilon$$

Since we know that:

$$\overline{\int_B} f \leq U(f, P)$$
$$\underline{\int_B} f \geq L(f, P)$$

Since $\varepsilon > 0$ is arbitrary, we can take $\varepsilon \to 0$ and so we must have that the upper and lower integrals agree. Therefore f is integrable.

With this we win! ☺

Proposition. *Let B be a box. Denote by $\mathcal{R}(B)$ the set of all Riemann integrbale functions on B. Then:*

1) $\mathcal{R}(B)$ is a vector space. That is if $f, g \in \mathcal{R}(B)$ then $f + cg \in \mathcal{R}(B)$ for all $c \in \mathbb{R}$. Furthermore, $\int_B$ is a linear function from $\mathcal{R}(B)$ to $\mathbb{R}$. That is:

$$\int_B f + cg = \int_B f + c \int_B g$$

2) Every constant function $f(x) = c$ is integrable, and in particular has integral $\int_B f = c \cdot v(B)$

3) If P is any partition of B then:

$$v(B) = \int_B 1 = \sum_Q v(Q)$$

Which is the sum taken over all sub-boxes determined by P

4) Let $B_1, \ldots, B_k$ be a collection of boxes that cover B, then:

$$v(B) \leq \sum_{j=1}^{k} v(B_j)$$

Proof. Let's go!

1) We leave this as an exercise

2 & 3) For any partition P note that:

$$L(f, P) = c \sum_{Q} v(Q) = U(f, P)$$

And therefore by the Riemann condition, f is integrable. And furthermore:

$$\int_B c = c \sum_{Q} v(Q)$$

Taking P to be the trivial partition we have that $\int_B c = c \cdot v(B)$

4) Let B be a box containing $B_1, \ldots, B_k$. Now let P be a partition of B that contains all the endpoints that define $B_1, \ldots, B_k$ and B. By the above:

$$v(B) = \sum_{Q \subseteq B} v(Q) \leq \sum_{j=1}^{k} \sum_{Q \subseteq B_i} v(Q) = \sum_{j=1}^{k} v(B_j)$$

MATH 395 Notes

Faye Jackson

November 4, 2020

Characterization of Riemann Integrability

Definition. *Let $A \subseteq \mathbb{R}^n$. We say that A has Lebesgue measure zero in $\mathbb{R}^n$ if for every $\varepsilon > 0$ there exists a covering of A by a countable collection $B_1, B_2, \dots$ of boxes such that:*

$$\sum_{j=1}^{\infty} v(B_j) < \varepsilon$$

We'll call this ℓ-measure zero for convenience.

Proposition. *Some properties of measure-zero sets:*

a) If $B \subseteq A$ and A has ℓ-measure zero, then B has ℓ-measure zero

b) If $A = \bigcup_{j=1}^{\infty} A_j$ and A_j has ℓ-measure zero for all j, then A has ℓ-measure zero.

c) A set A has ℓ-measure zero if and only if for every $\varepsilon > 0$ there exists a covering of A by a countable collection of open boxes $B_1, B_2, \dots$ such that:

$$\sum_{j=1}^{\infty} v(B_j) < \varepsilon$$

Aka, we may replace the boxes in the definition by open boxes

d) If B is a box, then ∂B has ℓ-measure zero

e) If $v(B) \neq 0$ then B does not have ℓ-measure zero

Proof. Let's go!

a) (a) is direct

b) Fix some $\varepsilon > 0$. Then since A_j has ℓ-measure zero there are boxes $B_{j1}, B_{j2}, \dots$ such that:

$$A_j \subseteq \bigcup_{k=1}^{\infty} B_{jk}$$
$$\sum_{k=1}^{\infty} v(B_{jk}) < \frac{\varepsilon}{2^j}$$

And then:

$$A \subseteq \bigcup_{j=1}^{\infty} \subseteq \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} B_{jk}$$
$$\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} v(B_{jk}) < \sum_{j=1}^{\infty} \frac{\varepsilon}{2^j} = \varepsilon$$

Therefore A has ℓ-measure zero.

c) The converse direction is immediate. We handle the forward direction. Let $A \subseteq \mathbb{R}^n$ have ℓ-measure zero. Fix $\varepsilon > 0$. We know that there is a collection of boxes $B_1, B_2 \dots$ such that:

$$A \subseteq \bigcup_{j=1}^{\infty} B_j$$
$$\sum_{j=1}^{\infty} v(B_j) < \frac{\varepsilon}{2^{n+1}}$$

Then for each B_j with $v(B_j) \neq 0$, let $\widetilde{B}_j$ be the open box that is obtained from B_j by dilating it (around its center), by a factor of 2. If $v(B_j) = 0$ then let $\widetilde{B}_j$

be an open box containing B_j with $v(\widetilde{B}_j) < \frac{\varepsilon}{2^{j+1}}$. Then clearly:

$$A \subseteq \bigcup_{j=1}^{\infty} B_j \subseteq \bigcup_{j=1}^{\infty} \widetilde{B}_j$$

$$\begin{aligned}\sum_{j=1}^{\infty} v(\widetilde{B}_j) &= \sum_{\substack{j=1 \\ v(B_j)=0}}^{\infty} v(\widetilde{B}_j) + 2^n \sum_{\substack{j=1 \\ v(B_j)\neq 0}}^{\infty} v(B_j) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\end{aligned}$$

Great! Thus A has ℓ-measure zero.

d) Let $B = [a_1, b_1] \times \cdots \times [a_n, b_n]$. Then ∂B is the union of the faces of B given by:

$$[a_1, b_1] \times [a_{j-1}, b_{j-1}] \times \xi_j \times [a_{j+1}, b_{j+1}] \times \cdots \times [a_n, b_n]$$

Where $1 \leq j \leq n$ and $\xi_j \in \{a_j, b_j\}$. Let us denote this ace by F_j. Then:

$$\begin{aligned} F_j \subseteq B_j &= [a_1, b_1] \times \cdots \times [\xi_j - \delta/2, \xi_j + \delta/2] \times \cdots \times [a_n, b_n] \\ v(B_j) &= \delta \prod_{i \neq j} b_j - a_j \end{aligned}$$

We can make this arbitrarily small by choosing δ to be small, and so F_j has ℓ-measure zero, showing that ∂B has ℓ-measure zero by part (b).

e) Now suppose that $v(B) \neq 0$ and B has ℓ-measure zero for the sake of ontradiction. We know that $\overline{B} = B \cup \partial B$ and so by part (b) we know that $\overline{B}$ has ℓ-measure zero, and also $v(\overline{B}) \neq 0$ since $B \subseteq \overline{B}$. Now take $\varepsilon = \frac{1}{2} v(\overline{B})$ and let $B_1, B_2, \ldots$ be a countable collection of open boxes such that:

$$\overline{B} \subseteq \bigcup_{i=1}^{\infty} B_j$$

$$\sum_{j=1}^{\infty} v(B_j) < \varepsilon$$

Since $\overline{B}$ is compact, there exists a finite subcollection, say $B_1, \ldots, B_k$ such taht

$\overline{B}$ is a subset of $B_1 \cup \cdots \cup B_k$. Then:

$$v(\overline{B}) \leq \sum_{j=1}^{k} v(B_j) < \varepsilon = \frac{1}{2} v(\overline{B})$$

Since $v(\overline{B}) \neq 0$ this gives a contradiction!

Great!

Example. The set of rational numbers in $[0,1]$ has ℓ-measure zero, because it is the countable union of singletons, and every singleton has ℓ-measure zero. Recall that this set is <u>not</u> Jordan measurable.

Theorem (Characterization of Riemann integrability)**.** *Let $B \subseteq \mathbb{R}^d$ be a box and $f : B \to \mathbb{R}$ be a bounded function. Let $\mathscr{D}$ be the set of points in B at which f is discontinuous. Then f is Riemann integrable on B if and only if $\mathscr{D}$ has ℓ-measure zero.*

Example. Consider the following function:

$$\begin{aligned} f(x) : [0,1] &\to \mathbb{R} \\ x &\overset{f}{\mapsto} \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases} \end{aligned}$$

Then $\mathscr{D} = [0,1]$, which does not have ℓ-measure zero. Therefore f is not Riemann integrable.

Proof. Choose M such that $|f(x)| \leq M$ for all $x \in B$:

($\Leftarrow$) Suppose that the set $\mathscr{D}$ has ℓ-measure zero. Let $\varepsilon > 0$ be given. We shall exhibit a partitioon P of B such that:

$$U(f,P) - L(f,P) \leq C\varepsilon$$

where C is a constant independent of ε and P. By the Riemann criterion, this implies that f is Riemann integrable. Since $\mathscr{D}$ has ℓ-measure zero. There exists open boxes $B_1, B_2, \dots$ such that:

$$\mathscr{D} \subseteq \bigcup_{j=1}^{\infty} B_j$$

$$\sum_{j=1}^{\infty} v(B_j) < \varepsilon$$

For each $x \notin \mathscr{D}$, f is continuous at x, and so there exists an open box Q_x centered at x such that:

$$\left|f(y) - f(y')\right| < \varepsilon \qquad (\forall y, y' \in Q_x \cap B)$$

Let $C_x = (Q_x \cap B)^o$ which is a box. The collection $\{B_j\}$ and $\{C_x\}$ is an open cover of B which is compact. Therefore there exists a finite subcover:

$$B_1 \cup \cdots \cup B_p \cup C_{x_1} \cup \cdots \cup C_{x_q}$$

Rename $C_\ell := C_{x_\ell}$. We have thus obtained that:

$$B = \left(\bigcup_{k=1}^{p} B_k\right) \cup \left(\bigcup_{\ell=1}^{q} C_\ell\right)$$

$$\sum_{k=1}^{p} v(B_k) < \varepsilon$$

$$y, y' \in C_\ell \implies \left|f(y) - f(y')\right| < \varepsilon$$

Let P be the partition of B that contains all of the endpoints of the component intervals of the boxes $\{B_k\}$ and $\{Q_\ell\}$. Then each B_k and each Q_ℓ is the union of sub-boxes is the union of sub-boxes determined by P.

We split the sub-boxes R determined by P into two groups, which we will call $\mathscr{R}_1$ and $\mathscr{R}_2$. $\mathscr{R}_1$ is the sub-boxes that are contained in B_k for some $1 \leq k \leq p$, then $\mathscr{R}_2$ are the sub-boxes contained in Q_ℓ for some $1 \leq \ell \leq q$.

We then estimate:

$$\begin{aligned}
U(f,P) - L(f,P) &= \sum_{R} [M_R(f) - m_R(f)] \cdot v(R) \\
&\leq \sum_{R \in \mathscr{R}_1} [M_R(f) - m_R(f)] \cdot v(R) + \sum_{R \in \mathscr{R}_2} [M_R(f) - m_R(f)] \cdot v(R) \\
&\leq \sum_{R \in \mathscr{R}_1} 2M \cdot v(R) + \sum_{R \in \mathscr{R}_2} \varepsilon \cdot v(R) \\
&\leq 2M \cdot \sum_{R \in \mathscr{R}_1} v(R) + \varepsilon \cdot \sum_{R \in \mathscr{R}_2} v(R) \\
&\leq 2M \cdot \sum_{k=1}^{p} \sum_{\substack{R \in \mathscr{R}_1 \\ R \subseteq B_k}} v(R) + \varepsilon \cdot \sum_{R} v(R) \\
&= 2M \cdot \sum_{k=1}^{p} v(B_k) + \varepsilon \cdot v(B) \\
&< (2M + v(B)) \cdot \varepsilon = C \cdot \varepsilon
\end{aligned}$$

And this finishes this part of the proof!

($\Rightarrow$) We now show that if f is integrable then $\mathscr{D}$ has ℓ-measure zero. We need to introduce the notion of the oscillation of a function at a point:

Definition. *With $g : A \subseteq \mathbb{R}^n \to \mathbb{R}$ bounded and for $x \in A$ we define the oscillation of g at x:*

$$\begin{aligned}
\operatorname{osc}_\delta g(x) &:= \sup_{y, y' \in A \cap B(x,\delta)} [g(y) - g(y')] \\
\operatorname{osc} g(x) &:= \inf_{\delta > 0} \operatorname{osc}_\delta g(x)
\end{aligned}$$

Exercise. *Show the following properties of the oscillation function:*

a) $\operatorname{osc}_\delta g(x) = \sup_{B(x,\delta) \cap A)} g - \inf_{B(x,\delta) \cap A} g \geq 0$.

b) $\operatorname{osc}_\delta g(x)$ *is increasing in* δ*, i.e. if* $\delta < \delta'$ *then* $\operatorname{osc}_\delta g(x) \leq \operatorname{osc}_{\delta'} g(x)$*. This follows because the supremum over a smaller set is smaller than the supremum over a bigger set*

c) Then we have that $\operatorname{osc} g(x) = \lim_{\delta \to 0} \operatorname{osc}_\delta g(x)$.

d) f *is continuous at* x *if and only if* $\operatorname{osc} f(x) = 0$.

The rest of this direction will be done in next section

MATH 395 Notes

Faye Jackson

November 9, 2020

Theorem. *Suppose $f : B \subseteq \mathbb{R}^n \to \mathbb{R}$ is bounded. Then f is Riemann integrable if and only if the set $\mathscr{D}$ of discontinuities of f has Lebesgue measure zero.*

Proof. We've already proved the $\impliedby$ direction in class. We are in the process of proving the $\implies$ direction using the properties of osc, which we defined at a point $x \in B$ as follows:

$$\begin{aligned} \operatorname{osc}_\delta f(x) &:= \sup_{y,y' \in B(x,\delta)\cap B} f(y) - f(y') \qquad (\delta > 0) \\ &:= \sup_{B(x,\delta)\cap B} f - \inf_{B(x,\delta)\cap B} f \\ \operatorname{osc} f(x) &:= \inf_{\delta>0} \operatorname{osc}_\delta f(x) = \lim_{\delta\to 0} \operatorname{osc}_\delta f(x) \end{aligned}$$

This holds because $\operatorname{osc}_\delta f(x)$ is increasing in δ.

Exercise. *Verify the properties of* osc *and* $\operatorname{osc}_\delta$*:*

a) $\operatorname{osc}_\delta f(x) = \sup_{B(x,\delta)} \cap B - \int_{B(x,\delta)} f \geq 0$

b) $\operatorname{osc}_\delta f$ *is increasing with* δ

c) f *is continuous at* $x \in B \iff \operatorname{osc} f(x) = 0$

Now we are ready to show that if f is Riemann integrable on B then $\mathscr{D}$ has Lebesgue measure zero:

$$\mathscr{D}_m := \left\{ x \in B \mid \operatorname{osc} f(x) \geq \frac{1}{m} \right\}$$

$$\mathscr{D} = \{ x \in B \mid \operatorname{osc} f(x) > 0 \} = \bigcup_{m=1}^{\infty} \mathscr{D}_m$$

Since $\mathscr{D}$ is a countable union of the $\mathscr{D}_m$, it suffices to show that each $\mathscr{D}_m$ has Lebesgue measure zero.

Let $\varepsilon > 0$ be arbitrary. We will cover $\mathscr{D}_m$ by countably many boxes whose total volume is less than ε. Note that since f is integrable we can find a partition P of B such that:

$$U(f,P) - L(f,P) < \frac{\varepsilon}{2m}$$

We now write that $\mathscr{D}_m = \mathscr{D}'_m \cup \mathscr{D}''_m$ where:

$$\begin{aligned}
\mathscr{D}'_m &= \{x \in \mathscr{D}_m \mid x \in \partial R \text{ for some sub-box } R \text{ determined by } P\} \\
\mathscr{D}''_m &= \mathscr{D}_m \setminus \mathscr{D}'_m
\end{aligned}$$

Note that $\mathscr{D}'_m \subseteq \bigcup_R \partial R$ where R ranges over the finitely many sub-boxes determined by P. Therefore, since we saw last time that the boundary of any box has Lebesgue measure zero, we know $\mathscr{D}'_m$ has Lebesgue measure zero. Of course we can then cover $\mathscr{D}'_m$ by countably many boxes whose total volume is less than $\frac{\varepsilon}{2}$

It remains to cover $\mathscr{D}''_m$ by countably many boxes of total volume less than $\frac{\varepsilon}{2}$. First note that if $x \in \mathscr{D}''_m$ then:

$$\begin{aligned}
&\operatorname{osc} f(x) \geq \frac{1}{2m} \\
&x \in R^\circ \text{ for some sub-box } R \text{ determined by the partition}
\end{aligned}$$

Therefore there exists a $\delta > 0$ so that $B(x,\delta) \subseteq R$ and:

$$\begin{aligned}
\frac{1}{2m} \leq \operatorname{osc} f(x) \leq \operatorname{osc}_\delta f(x) &= \sup_{B(x,\delta)} f - \inf_{B(x,\delta)} f \\
&\leq \sup_R f - \inf_R f = M_R(f) - m_R(f)
\end{aligned}$$

We multiply by $v(R)$ and summing over all R we get:

$$\begin{aligned}
\frac{1}{2m} \sum_{\substack{R \\ R\cap \mathscr{D}''_m \neq \emptyset}} v(R) &\leq \sum_{\substack{R \\ R\cap \mathscr{D}''_m \neq \emptyset}} (M_R(f) - m_R(f)) \cdot v(R) \\
&\leq \sum_{R} (M_R(f) - m_R(f)) \cdot v(R) \\
&= U(f,P) - L(f,P) < \frac{\varepsilon}{2m}
\end{aligned}$$

And therefore:

$$\sum_{\substack{R \\ R\cap \mathscr{D}''_m \neq \emptyset}} < \frac{\varepsilon}{2}$$

These boxes which intersect $\mathscr{D}''_m$ provide the needed covering of $\mathscr{D}''_m$.

Remark. This theorem shows that sets of Lebesgue measure zero can be problematic for Riemann integration. In the sense that, changing a function on a set of Lebesgue measure zero can make it non-integrable. In particular consider the function:

$$\begin{aligned}
\mathbb{1}_{\mathbb{Q}} &: [0,1] \to \mathbb{R} \\
\mathbb{1}_{\mathbb{Q}}(x) &= \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}
\end{aligned}$$

This is only different from a constant function on a set of measure zero, namely it differs from the constant function on $\mathbb{Q} \cap [0,1]$. This indicates a kind of "incompleteness" of Riemann integration.

Corrolary. *Let B be a box in $\mathbb{R}^n$ and $f : B \to \mathbb{R}$ be Riemann integrable.*

a) If f vanishes except on a set of Lebesgue measure zero, then $\int_B f = 0$. We say that $f = 0$ almost everywhere

b) If $f \geq 0$ and $\int_B f = 0$ then f vanishes except possibly on a set of Lebesgue measure zero. That is f vanishes almost everywhere.

Remark. The corollary is not true without the assumption that f is Riemann integrable.

Proof. Let's go!

a) Let $\mathscr{D}_0$ be the set $\{x \in B \mid f(x) \neq 0\}$. By assumption, $\mathscr{D}_0$ has ℓ-measure zero. Let P be any partition of B. For any sub-box R of this partition, we have that $R \nsubseteq \mathscr{D}_0$ (since $v(R) > 0$). This implies there exists an $x \in R$ such that $f(x) = 0$, and so:

$$m_R(f) \leq 0$$
$$M_R(f) \geq 0$$

Therefore $L(f, P) \leq 0$ and $U(f, P) \geq 0$. But wait this implies that:

$$\underline{\int_B} f \leq 0 \qquad\qquad \overline{\int_B} f \geq 0$$

Since f is integrable, we then know that:

$$\int_B f = \underline{\int_B} f = \overline{\int_B} f = 0$$

And so we are done.

b) Suppose $f(x) \geq 0$ and $\int_B f = 0$. We will show that if f is continuous at some x, then $f(x) = 0$. Since the set of discontinuities of f has measure zero beause f is Riemann integrable, this shows that the set of all x where $f(x) \neq 0$ must have measure zero as well.

We will do this by contradiction. Suppose that f is continuous at some x_0 and $f(x_0) > 0$. Then there exists an $\varepsilon > 0$ and a small box R centered at x_0 such that $f(x) > \varepsilon$ for all $x \in R$.

Now consider the following function:

$$g(x) = \begin{cases} \varepsilon & \text{if } x \in \mathbb{R} \\ 0 & \text{if } x \in B \setminus R \end{cases}$$

Then g is integrable since the set of discontinuities of g has measure zero. Also $f(X) \geq g(x)$ for all $x \in B$ and so:

$$\int_B f \overset{?}{\geq} \int_B g \overset{?}{=} \varepsilon \cdot v(R) > 0$$

Hani says we should verify $\overset{?}{\geq}$ and $\overset{?}{=}$. I leave that to you ☺

Another approach is to take a partition P obtained from the endpoints of R and B. Then we get $L(f,P) \geq \varepsilon \cdot v(R_0)$. where R_0 is the sub-box of P containing x_0. But this implies that:

$$\int_B f(x)\,\mathrm{d}x = \sup_P L(f,P) \geq \varepsilon \cdot v(R_0)$$

In either case, we have an oops! Great!

Fubini's Theorem

After defining the integral, the main question remains: how to compute integrals in higher dimensions? (We know how to compute integrals in 1D using the Fundamental Theorem of Calculus and various techniques of integration)

Fubini's Theorem will allow us to compute integrals in higher dimensions by reducing them to iterated integrals in lower dimensions. This often allows us to reduce things to the one-dimensional case.

One would wish to say that if $f : Q \to \mathbb{R}$ is integrable where $Q = A \times B$ and A is a box in $\mathbb{R}^k$ and B is a box in $\mathbb{R}^\ell$. Then $x \mapsto \int_B f(x,y)\,\mathrm{d}y$ exists for every $x \in A$ and defines an integrable function over A. Furthermore:

$$\int_Q f = \int_A \left(\int_B f(x,y)\,\mathrm{d}y \right) \mathrm{d}x \tag{$\star$}$$

This requires that the function $x \mapsto \int_B f(x,y)\,\mathrm{d}y$ is defined for every x (i.e. $f(x,y)$ is integrable in y for fixed $x \in A$) and that function $x \mapsto \int_B f(x,y)\,\mathrm{d}y$ is integrable in x itself on A.

Unfortunately, such a nice property is not necessarily true for all $x \in A$. Indeed, we will see that it is true except for sets of Lebesgue measure zero. This is no problem for Lebesgue integrals (for which $\star$ holds), but since Riemann integrability can depend on sets of Lebesgue measure zero, we might lose there.

Theorem (Fubini's Theorem)**.** *Let $Q = A \times B$ where A is a box in $\mathbb{R}^k$ and B is a box in $\mathbb{R}^\ell$. let $f(x,y) : Q \to \mathbb{R}$ be a bounded function (where $x \in A$ and $y \in B$)*

Then for each $x \in A$ consider the lower and upper integrals:

$$x \mapsto \underline{\int_B} f(x,y)\,\mathrm{d}y \qquad\qquad x \mapsto \overline{\int_B} f(x,y)\,\mathrm{d}y$$

if f is integral over Q then the above two functions are integrable over A and:

$$\int_Q f = \int_A \left(\underline{\int_B} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x = \int_A \left(\overline{\int_B} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$

Of course we have lower and upper integrals here. If we get agreement of the above two functions on all of x then we would be very happy.

MATH 395 Notes

Faye Jackson

November 11, 2020

Theorem (Fubini)**.** *Given a box $Q = A \times B$ where $A \subseteq \mathbb{R}^k$ and $B \subseteq \mathbb{R}^\ell$ are boxes. Let $f : Q \to \mathbb{R}$ be a bounded function, and we write it as $f(x, y)$ where $x \in A$ and $y \in B$.*

If f is integrable over Q, then the functions:

$$x \mapsto \underline{\int_B} f(x, y)\, \mathrm{d}y \qquad\qquad x \mapsto \overline{\int_B} f(x, y)\, \mathrm{d}y$$

are both integrable over A, and furthermore:

$$\int_Q f = \int_A \underline{\int_B} f(x, y)\, \mathrm{d}y\, \mathrm{d}x = \int_A \overline{\int_B} f(x, y)\, \mathrm{d}y\, \mathrm{d}x$$

Remark. The drawback is that the iterated integrals are in terms of lower and upper integrals rather than having $\int_B f(x, y)\, \mathrm{d}y$. We cannot guarantee that these agree

Corrolary. *With the same assumptions as above, there holds the following:*

a) *$\int_B f(x, y)\, \mathrm{d}y$ exists for almost every $x \in A$, that is, it exists except on a set of Lebesgue measure zero. In other words $x \mapsto \int_B f(x, y)\, \mathrm{d}y$ is defined for all $x \in A \setminus N$ where N has Lebesgue measure zero.*

b) *If we further assume that $\int_B f(x, y)\, \mathrm{d}y$ exists for all $x \in A$, then we have Fubini's Theorem as we would like it:*

$$\int_Q f = \int_A \int_B f(x, y)\, \mathrm{d}y\, \mathrm{d}x$$

c) *Let $Q = I_1 \times \cdots \times I_n$ where $I_k = [a_i, b_i] \subseteq \mathbb{R}$. Then if $f : Q \to \mathbb{R}$ is continuous* *then:*

$$\int_Q f = \int_{I_1} \cdots \int_{I_n} f(x_1, \ldots, x_n) \, \mathrm{d}x_n \cdots \mathrm{d}x_1$$

Proof of Corollary. By Fubini's:

$$\int_Q f = \int_A \underline{\int_B} f(x, y) \, \mathrm{d}y \, \mathrm{d}x = \int_A \overline{\int_B} f(x, y) \, \mathrm{d}y \, \mathrm{d}x$$

$$0 = \int_A \left(\underbrace{\overline{\int_B} f(x, y) \, \mathrm{d}y - \underline{\int_B} f(x, y) \, \mathrm{d}y}_{\geq 0 \text{ and integrable}} \right) \mathrm{d}x$$

Therefore by previous work:

$$\overline{\int_B} f(x, y) \, \mathrm{d}y - \underline{\int_B} f(x, y) = 0$$

except possibly on a set of measure zero. This gives part a).

Part b) is exactly from Fubini's theorem, and part c) follows because continuous functions are always integrable.

Proof of Fubini. Let us define the following:

$$\underline{I}(x) = \underline{\int_B} f(x, y) \, \mathrm{d}y \qquad \overline{I}(x) = \overline{\int_B} f(x, y) \, \mathrm{d}y$$

We need to show that if $\int_Q f$ exists then $\underline{I}(x)$ and $\overline{I}(x)$ are both integrable over A, and their integrals are both $\int_Q f$.

Let P be any partition of Q and write $P = (P_A, P_B)$ are partitions of A and B. Any sub-box R determined by P can be written as $R = R_A \times R_B$ where R_A and R_B are sub-boxes of A and B determined by P_A and P_B respectively.

Now note that for any $x \in R_A$:

$$m_R(f) = \inf_R f(x, y) \leq \inf_{y \in R_B} f(x, y)$$

$$m_R(f) \leq m_{R_B}(f(x, -))$$

Multiplying by $v(R_B)$ and summing over all sub-boxes R_B we get for every $x \in R_A$:

$$\begin{aligned}\sum_{R_B} m_R(f) \cdot v(R_B) &\leq \sum_{R_B} m_{R_B}(f(x,-))v(R_B) \\ &= L(f(x,-), P_B) \leq \underline{I}(x)\end{aligned}$$

Then if we take the infimum over $x \in R_A$ we obtain:

$$\sum_{R_B} m_R(f)v(R_B) \leq m_{R_A}(\underline{I})$$

We then multiply by $v(R_A)$ and sum over all such R_A:

$$\begin{aligned}\sum_{R_A, R_B} m_R(f)v(R_B)v(R_A) &\leq \sum_{R_A} m_{R_A}(\underline{I})v(R_A) \\ L(f, P) &\leq L(\underline{I}, P_A)\end{aligned}$$

An exactly similar argument establishes that:

$$U(f, P) \geq U(\overline{I}, P_A)$$

Given these two inequalities, we will be able to finish the proof. Note that because $\underline{I} \leq \overline{I}$ we have:

$$\begin{aligned}L(f, P) \leq L(\underline{I}, P_A) \leq U(\underline{I}, P_A) \leq U(\overline{I}, P_A) \leq U(f, P) \\ L(f, P) \leq L(\underline{I}, P_A) \leq L(\overline{I}, P_A) \leq U(\overline{I}, P_A) \leq U(f, P)\end{aligned}$$

These inqualities hold for any partition P. Let $\varepsilon > 0$ be arbitrary and choose P so that $U(f, P) - L(f, P) < \varepsilon$. Therefore from the above inequalities and a squeezing argument:

$$\begin{aligned}U(\underline{I}, P_A) - L(\underline{I}, P_A) < \varepsilon \\ U(\overline{I}, P_A) - L(\overline{I}, P_A) < \varepsilon\end{aligned}$$

This gives that $\underline{I}$ and $\overline{I}$ are both integrable on A. Now we get that:

$$L(f,P) \leq \int_Q f \leq U(f,P)$$
$$L(f,P) \leq L(\underline{I},P) \leq \int_A \underline{I} \leq U(\underline{I},P) \leq U(f,P)$$
$$L(f,P) \leq L(\overline{I},P) \leq \int_A \overline{I} \leq U(\overline{I},P) \leq U(f,P)$$

Therefore we get that:

$$\left|\int_Q f - \int_A \underline{I}\right| < \varepsilon \qquad\qquad \left|\int_Q f - \int_A \overline{I}\right| < \varepsilon$$

And so since $\varepsilon > 0$ was chosen arbitrarily, we must have that:

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

Which is exactly what we wanted to show!

Integral over a bounded set

Up until now, we have been integrating functions on boxes. What if we want to integrate a function over a region $S \subseteq \mathbb{R}^n$ that is not a box.

Definition. *Let $S \subseteq \mathbb{R}^n$ be a bounded set and suppose that $f : S \to \mathbb{R}$ is a bounded function. We define $f_S(x) = f(x)$ when $x \in S$ and $f_S(x) = 0$ when $x \notin S$. Then f_S is defined on all of $\mathbb{R}^n$, that is $f_S : \mathbb{R}^n \to \mathbb{R}$*

Choose a box Q which contains S, then we define the integral of f over S as:

$$\int_S f(x)\,\mathrm{d}x = \int_Q f_S(x)\,\mathrm{d}x$$

provided that the integral on the right hand side exists.

For this definition to make sense, we should get the same answer if we change the box Q. This is guaranteed by the following lemma:

Lemma. *Let Q and Q' be two boxes in $\mathbb{R}^n$ and let $f : \mathbb{R}^n \to \mathbb{R}$ be a function that is supported inside $Q \cap Q'$. That is $f = 0$ outside $Q \cap Q'$.*

$$\int_Q f = \int_{Q'} f$$

Included in the statement is that f is integrable over Q if and only if f is integrable over Q'.

Proof. Let's go!

Case 1) Suppose that $Q \subseteq Q'$. Then f is supported in Q.

Note that f is integrable on Q if and only if the set of discontinuities of f in Q has Lebesgue measure zero, we call this set $\mathscr{D}$. But wait! The set of discontinuities of f in Q', which we call $\mathscr{D}'$, is equal to $\mathscr{D} \cup A$, where $A \subseteq \partial Q$, because f is constant on $Q' \setminus Q$. Since ∂Q has Lebesgue measure zero, and so A has Lebesgue measure zero, we know $\mathscr{D}'$ has Lebesgue measure zero if and only if $\mathscr{D}$ has Lebesgue measure zero. Therefore:

$$f \text{ is integrable over } Q' \iff f \text{ is integrable over } Q$$

Now let P be a partition of Q' and let $\widetilde{P}$ be the refinement of P obtained from P by adding in the endpoints of Q. Then Q is the union of some sub-boxes determined by $\widetilde{P}$. Write $Q = \bigcup_{B \in \mathscr{S}}$ where $\mathscr{S}$ is the family of sub-boxes determined by $\widetilde{P}$ such that $B \subseteq Q$.

Now if B is determined by $\widetilde{P}$ and $B \notin \mathscr{S}$ then $f(x) = 0$ for some $x \in B$.

Therefor if $B \notin \mathscr{S}$ then $m_B(f) \leq 0 \leq M_B(f)$. Therefore:

$$L(f, P) \leq L(f, \widetilde{P}) \leq \sum_{B \in \mathscr{S}} m_B(f) v(B) \leq \int_Q f$$
$$U(f, P) \geq U(f, \widetilde{P}) \geq \sum_{B \in \mathscr{S}} M_B(f) v(B) \geq \int_Q f$$

This holds for any P. Taking suprema and infima in P:

$$\int_{Q'} f = \sup L(f, P) \leq \int_Q f$$
$$\int_{Q'} f = \inf U(f, P) \geq \int_Q f$$

And therefore $\int_{Q'} f = \int_Q f$

Case 2) Pick Q'' to be a sufficiently large box containing both Q and Q'. Then:

$$\int_Q f = \int_{Q''} f = \int_{Q'} f$$

Just by applying Case 1 twice, and of course existence of these integrals if and only if one of them exists.

Handout 9

Jordan measure and Riemann Integration

It turns out that the notion of Jordon measurability of sets is intimately related (in a way essentially equivalent) to the notion of Riemann integrability of functions. We will only display this relation in dimension 1.

- **Recall.** To define the Riemann[1] integral of a <u>bounded</u> function f on an interval $[a, b] \subset \mathbb{R}$, we first recall the notion of a partition $\mathcal{P}$ which is a set of points $x_0 = a < x_1 < x_2 < \ldots < x_n = b$, the norm of the partition is $\Delta\mathcal{P} = \max_{1\leq k\leq n} x_k - x_{k-1}$, and we denote by $\Delta x_k = x_k - x_{k-1}$. For each such partition, we define to quantities:

$$L(f, \mathcal{P}) = \sum_{k=1}^{n} f(x_*)\Delta x_k, \quad and \quad U(f, \mathcal{P}) = \sum_{k=1}^{n} f(x^*)\Delta x_k,$$

where $x_* = \inf_{[x_{k-1}, x_k]} f$ and $x^* = \sup_{[x_{k-1}, x_k]} f$.

Afterwards, we define the lower and upper Darboux integrals respectively as

$$\underline{\int_a^b} f(x)dx = \sup_{\mathcal{P}} L(f, \mathcal{P}), \quad and \quad \overline{\int_a^b} f(x)dx = \inf_{\mathcal{P}} U(f, \mathcal{P}).$$

where the extrema above are taken over all partitions of the interval $[a, b]$. We say that f is Riemann integrable if the above two numbers are equal. We define the common value as the Riemann (or Darboux) integral of f.

[1] Strictly speaking, we are recalling here the notion of Darboux integral, but that is equivalent to the notion of Riemann integrability that is often covered in introductory calculus classes.

Q1) Let $[a, b]$ be an interval and let $f : [a, b] \to \mathbb{R}$ be a bounded nonnegative function. Show that f is Riemann integrable if and only if the set $E := \{(x, t) : x \in [a, b] : 0 \leq t \leq f(x)\}$ is Jordan measurable in $\mathbb{R}^2$.

Q2) Let $[a, b]$ be an interval and let $f : [a, b] \to \mathbb{R}$ be a bounded function. Show that f is Riemann integrable if and only if the sets $E_+ := \{(x, t) : x \in [a, b] : 0 \leq t \leq f(x)\}$ and $E_- := \{(x, t) : x \in [a, b] : f(x) \leq t \leq 0\}$ are Jordan measurable in $\mathbb{R}^2$.

Remark. The above results generalize to higher dimensions.

Where we are right now?

We have thus far discussed the classical theory of Jordan measure, which went as follows

(i) We define the notion of a box and its volume $|B|$ or $v(B)$,

(ii) Then we defined the notion of an elementary set and its elementary measure,

(iii) Then we defined the notion of Jordan inner and outer measure $\underline{m}_J(E)$ and $\overline{m}^J(E)$ and said that a set E is Jordan measurable if those two concepts agree.

In particular, unwinding the definition of the Jordan outer measure, we have that for any set E

$$\overline{m}_J(E) = \inf_{E \subset B_1 \cup \ldots \cup B_k} |B_1| + \ldots + |B_k|$$

where the infimum is taken over all finite coverings of E by boxes $B_1, \ldots, B_k$.

Q3) Show that a set E is Jordan measurable if and only if for every $\epsilon > 0$ there exists an elementary set U containing E such that $\overline{m}_J(U \setminus E) < \epsilon$.

The notions of Lebesgue outer measure and Lebesgue measurability are refinements of the Jordan ones as follows:

– **Lebesgue outer measure:** We modify the notion of Jordan outer measure by replacing the finite union of boxes by a countable union of boxes, i.e.

$$m^*(E) = \inf_{E \subset \cup_{j=1}^{\infty} B_j} \sum_{j=1}^{\infty} |B_j|$$

where the union above is taken over boxes $B_j \subset \mathbb{R}^d$.

Q4) Show that the Lebesgue outer measure $m^*(E)$ is zero for any countable set E. Contrast this to fact that the Jordan outer measure of the rationals in $[0, 1]$ was equal to 1.

– **Lebesgue measurability** A set $E \subset \mathbb{R}^d$ is said to be Lebesgue measurable if for every $\epsilon > 0$, there exists an open set $U \subset \mathbb{R}^d$ containing E such that $m^*(U \setminus E) \leq \epsilon$. If E is measurable, we refer to $m(E) = m^*(E)$ as the Lebesgue measure of E.

Remarks: Note that there is no need for E to be bounded for this definition to make sense. Also, the notion of Lebesgue measurability can be seen as a (finite to countably infinite) generalization of that of Jordan measurability since it can be shown that every open set is the countable union of closed boxes.

Handout 10

Where we are right now?

- We have thus far discussed the classical theory of Jordan measure, which went as follows
  (i) We define the notion of a box and its volume $|B|$ or $v(B)$,
  (ii) Then we defined the notion of an elementary set and its elementary measure,
  (iii) Then we defined the notion of Jordan inner and outer measure $\underline{m}_J(E)$ and $\overline{m}^J(E)$ and said that a set E is Jordan measurable if those two concepts agree.

  In particular, unwinding the definition of the Jordan outer measure, we have that for any set E

  $$\overline{m}_J(E) = \inf_{E \subset B_1 \cup \ldots \cup B_k} |B_1| + \ldots + |B_k|$$

  where the infimum is taken over all finite coverings of E by boxes $B_1, \ldots, B_k$.

Q0) Show that a set E is Jordan measurable if and only if for every $\epsilon > 0$ there exists an elementary set U containing E such that $\overline{m}_J(U \setminus E) < \epsilon$.

Lebesgue outer measure

The notions of Lebesgue outer measure and Lebesgue measurability are refinements of the Jordan ones as follows:

- **Lebesgue outer measure:** We modify the notion of Jordan outer measure by replacing the finite union of boxes by a countable union of boxes, i.e.

  $$m^*(E) = \inf_{E \subset \cup_{j=1}^{\infty} B_j} \sum_{j=1}^{\infty} |B_j|$$

where the union above is taken over boxes $B_j \subset \mathbb{R}^d$.

Q1) Show that $m^*(E) \leq \overline{m}_J(E)$ where $\bar{m}_J$ is the Jordan outer measure.

Q2) Show that in the definition above the countable cover by boxes in the definition of $m^*(E)$ can be restricted to closed boxes or open boxes.

Q3) Show that the Lebesgue outer measure $m^*(E)$ is zero for any countable set E. Contrast this to fact that the Jordan outer measure of the rationals in $[0, 1]$ was equal to 1.

- **Lebesgue measurability** A set $E \subset \mathbb{R}^d$ is said to be Lebesgue measurable if for every $\epsilon > 0$, there exists an open set $U \subset \mathbb{R}^d$ containing E such that $m^*(U \setminus E) \leq \epsilon$. If E is measurable, we refer to $m(E) = m^*(E)$ as the Lebesgue measure of E.

  Remarks:

  (i) Note that there is no need for E to be bounded for this definition to make sense.

  (ii) The notion of Lebesgue measurability can be seen as a (finite to countably infinite) generalization of that of Jordan measurability since it can be shown that every open set is the countable union of closed boxes.

Q4) Show that $m^*(\emptyset) = 0$.

Q5) (Monotonicity) Show that if $E \subset F \subset \mathbb{R}^d$, then $m^*(E) \leq m^*(F)$.

Q6) (Countable subadditivity) If $E_1, E_2, \ldots \subset \mathbb{R}^d$ is a countable sequence of sets, then $m^*\left(\cup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} m^*(E_n)$.

MATH 395 Notes

Faye Jackson

November 13, 2020

Exercise 0. *Show that a set E is Jordan measurable if and only if for every $\varepsilon > 0$ there exists an elementary set U containing E such that $\overline{m}_J(U \setminus E) < \varepsilon$.*

Proof. **TODO**

Exercise 1. *Show that $m^*(E) \leq \overline{m}_J(E)$ where $\overline{m}_J$ is the Jordan outer measure*

Proof. Fix some elementary set A which contains E and write it as the disjoint union of a finite collection of boxes $B_1, \ldots, B_n$ that cover E. Then note that:

$$m^*(E) = \inf_{E \subseteq \bigcup_{j=1}^{\infty} C_j} \sum_{j=1}^{\infty} |C_j| \leq \sum_{j=1}^{n} |B_j| = m(A)$$

And so taking the infimum over all elementary sets A containing E we obtain:

$$m^*(E) \leq \overline{m}_J(E)$$

Just as desired.

Exercise 2. *Show that in the definition above the countable cover by boxes in the definition of $m^*(E)$ can be restriced to closed boxes or open boxes*

Proof. We deal with closed boxes first. Consider the sets:

$$S = \left\{ \sum_{j=1}^{\infty} |B_j| \mid E \subseteq \bigcup_{j=1}^{\infty} B_j \right\}$$

$$S_c = \left\{ \sum_{j=1}^{\infty} |B_j| \mid E \subseteq \bigcup_{j=1}^{\infty} B_j, B_j \text{ closed} \right\}$$

$$S_o = \left\{ \sum_{j=1}^{\infty} |B_j| \mid E \subseteq \bigcup_{j=1}^{\infty} B_j, B_j \text{ open} \right\}$$

We know that $m^*(E) = \inf S$ and we wish to show that $\inf S = \inf S_c = \inf S_o$. Now note that of course $S_o, S_c \subseteq S$, and so $\inf S \leq \inf S_c, \inf S_o$, therefore it only remains to show that $\inf S \geq \inf S_c, \inf S_o$.

To do so, by definition of greatest lower bound, it suffices to show that $\inf S_c$ and $\inf S_o$ are both lower bounds for S. We handle each of these:

- Take some countable collection of boxes $B_1, B_2, \ldots$ such that their union contains E, giving us an element $\sum |B_j|$ of S. Then we may consider the collection of their closures $\overline{B}_1, \overline{B}_2, \ldots$. Since $B_j \subseteq \overline{B}_j$ we know that the union of all these contains E. So then $\sum \left|\overline{B}_j\right| \in S_c$. But then we are in a great spot! We know $\left|\overline{B}_j\right| = |B_j|$. So then we may write:

  $$\inf S_c \leq \sum_{j=1}^{\infty} \left|\overline{B_j}\right| \leq \sum_{j=1}^{\infty} |B_j|$$

  And so $\inf S_c$ is a lower bound for S, and so $\inf S_c \leq \inf S$ as desired.

- Take some countable collection of boxes $B_1, B_2, \ldots$ whose union contains E, giving us an element $\sum |B_j|$ of S. We will show for any $\varepsilon > 0$ that:

  $$\inf S_o \leq \varepsilon + \sum_{j=1}^{\infty} |B_j|$$

  And so taking $\varepsilon \to 0$ we see that $\inf S_o$ is a lower bound for S and so $\inf S_o \leq \inf S$ as desired.

  Fix some such $\varepsilon > 0$, and consider the open box C_j obtained from B_j by dilating B_j so that $|C_j| \leq |B_j| + \frac{\varepsilon}{2^j}$ and $B_j \subseteq C_j$. Then $\sum C_j$ lies in S_o since the union of all the C_j contains E. But then:

$$\inf S_o \leq \sum_{j=1}^{\infty} |C_j| \leq \sum_{j=1}^{\infty} \frac{\varepsilon}{2^j} + \sum_{j=1}^{\infty} |B_j| = \varepsilon + \sum_{j=1}^{\infty} |B_j|$$

Taking $\varepsilon \to 0$ we see that $\inf S_o \leq \sum |B_j|$, and so $\inf S_o$ is a lower bound for S, giving us that $\inf S_o \leq \inf S$ as desired.

With this we are done! $m^*(E) = \inf S = \inf S_o = \inf S_c$. Great!

Exercise 3. *Show that the Lebesgue outer measure $m^*(E)$ is zero for any countable set E. Cotnrast this to the face that the Jordan outer measure of the rationals in $[0,1]$ was equal to 1*

Proof. Let E be a countable set. Then consider that:

$$E \subseteq \bigcup_{x \in E} \{x\}$$

exhibits E as a countable union of boxes, all of measure zero. Therefore:

$$0 \leq m^*(E) \leq \sum_{x \in E} |\{x\}| = 0$$

Showing us that $m^*(E) = 0$.

Let's look for another way of doing this! Write E as $x_1, x_2, \ldots$. We will allow repeats here, and if E is empty just repeat $x_n = 0$. Fix $\varepsilon >)$ and then take the box whose volume is $\frac{\varepsilon}{2^j}$ around every point $x_j = (x_{j1}, \ldots, x_{jd})$. In other words:

$$\begin{aligned}
B_j &= \prod_{k=1}^{d} \left[x_{jk} - \frac{\sqrt[d]{\varepsilon}}{2\sqrt[d]{2^j}}, x_{jk} + \frac{\sqrt[d]{\varepsilon}}{2\sqrt[d]{2^j}} \right] \\
|B_j| &= \prod_{k=1}^{d} \frac{\sqrt[d]{\varepsilon}}{\sqrt[d]{2^j}} = \frac{\varepsilon}{2^j} \\
\sum_{j=1}^{\infty} |B_j| &= \sum_{j=1}^{\infty} \frac{\varepsilon}{2^j} = \varepsilon
\end{aligned}$$

Great! Since $E \subseteq \bigcup_{j=1}^{\infty} B_j$ this means that:

$$0 \leq m^*(E) \leq \sum_{j=1}^{\infty} |B_j| = \varepsilon$$

Now taking $\varepsilon \to 0$ we get $m^*(E) = 0$.

Exercise 4. *Show that $m^*(\emptyset) = 0$.*

Proof. Note that $\emptyset$ is a countable set, so this follows easily from Q3

Exercise 5. *Show that if $E \subseteq F \subseteq \mathbb{R}^d$ then $m^*(E) \leq m^*(F)$.*

Proof. We will show that $m^*(E)$ is a lower bound for the set defining $m^*(F)$, and so by definition of infimum we have $m^*(E) \leq m^*(F)$.

Fix some countable collection of boxes $B_1, B_2, \ldots$ containing F, then in particular they contain E since F contains E, and so by definition of infimum:

$$m^*(E) \leq \sum_{j=1}^{\infty} |B_j|$$

Taking the infimum on the right hand side we get:

$$m^*(E) \leq m^*(F)$$

Great! This is exactly what we want!

Exercise 6. *If $E_1, E_2, \ldots \subseteq \mathbb{R}^d$ is a countable sequence of sets, then:*

$$m^*\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} m^*(E_n)$$

Great!

Proof. Fix some $\varepsilon > 0$, we will show that:

$$m^*\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \varepsilon + \sum_{n=1}^{\infty} m^*(E_n)$$

and so by taking $\varepsilon \to 0$ we will obtain the result. Take $E = \bigcup_{n=1}^{\infty} E_n$ for convenience.

Consider some E_n, then by definition of infimum and the fact that $\frac{\varepsilon}{2^n} > 0$ there is some countable collection of boxes $B_{n1}, B_{n2}, \ldots$ containing E_n such that:

$$m^*(E_n) \leq \sum_{j=1}^{\infty} |B_{nj}| \leq m^*(E_n) + \frac{\varepsilon}{2^n}$$

We can then sum over all E_n to get:

$$\sum_{n=1}^{\infty} \sum_{j=1}^{\infty} |B_{nj}| \leq \sum_{n=1}^{\infty} m^*(E_n) + \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon + \sum_{n=1}^{\infty} m^*(E_n)$$

And so now consider the countable collection of all the $\{B_{nj}\}$. This will be countable by 295, and also it will cover E, since for every $x \in E$ we know $x \in E_n$ for some n and then by construction $x \in B_{nj}$ for some j. But then by definition of infimum:

$$m^*(E) \leq \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} |B_{nj}| \leq \varepsilon + \sum_{n=1}^{\infty} m^*(E_n)$$

Since $\varepsilon > 0$ was chosen to be arbitrary, we can take $\varepsilon \to 0$ and we see that:

$$m^*\left(\bigcup_{n=1}^{\infty} E_n\right) = m^*(E) \leq \sum_{n=1}^{\infty} m^*(E_n)$$

Great! This is the desired result ☺.

MATH 395 Notes

Faye Jackson

November 18, 2020

Continuing the characterization of Jordan Measurability

Theorem. *Let S be a bounded subset of $\mathbb{R}^n$. The following are equivalent:*

1) S is Jordan measurable

2) The constant function 1 is Riemann Integrable on S

3) ∂S has Lebesgue measure zero

4) ∂S has Jordan outer measure zero.

Proof. Let's go!

$1 \implies 2)$ Suppose S is Jordan measurable. We need to show that:

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

is Riemann integrable on some box B containing S. Now let $\varepsilon > 0$ be arbitrary and pick two elementary sets $E_1 \subseteq S \subseteq E_2$ such that $m(E_2 \setminus E_1) < \varepsilon$. Without loss of generality, by dilating the component boxes of E_2 we may assume that $S \subseteq E_2^\circ$.

Choose B to be some box containing E_2. Now let P be a partition B that contains the endpoints of the intervals defining the boxes whose union is E_1 and E_2. Let $R_1, \ldots, R_m$ be some enumeration of the sub-boxes determined by this partition. Then:

$$\begin{aligned} U(\mathbb{1}_S, P) &= \sum_{i=1}^{m} M_{R_i}(\mathbb{1}_S) v(R_i) \\ &= \sum_{R_i \cap S \neq \emptyset} M_{R_i}(\mathbb{1}_S) v(R_i) \\ &\leq \sum_{R_i \subseteq E_2} M_{R_i}(\mathbb{1}_S) v(R_i) \\ &\leq \sum_{R_i \subseteq E_2} v(R_i) = m(E_2) \end{aligned}$$

Similarly, we can show that $L(\mathbb{1}_S, P) \geq m(E_1)$. But then:

$$U(\mathbb{1}_S, P) - L(\mathbb{1}_S, P) \leq m(E_2) - m(E_1) = m(E_2 \setminus E_1) < \varepsilon$$

Great! Therefore $\mathbb{1}_S$ is integrable and:

$$m(E_1) \leq L(\mathbb{1}_S, P) \leq \int_S 1 \, \mathrm{d}x \leq U(\mathbb{1}_S, P) \leq m(E_2)$$

and:

$$m(E_1) \leq m(S) \leq m(E_2)$$

Gives us that:

$$\left| \int_S 1 \, \mathrm{d}x - m(S) \right| < \varepsilon$$

For any $\varepsilon > 0$, and therefore:

$$m(S) = \int_S 1 \, \mathrm{d}x$$

$2 \implies 1)$ Let B be a box which contains S and take $\varepsilon > 0$ to be arbitrary. Since $\mathbb{1}_S$ is integrable on B, there exists a partition P of B such that:

$$U(\mathbb{1}_S, P) - L(\mathbb{1}_S, P) < \varepsilon$$

Let $R_1, \ldots, R_m$ be an enumeration of the sub-boxes determined by P. Now set:

$$E_1 = \bigcup_{R_i \subseteq S} R_i \subseteq S$$
$$E_2 = \bigcup_{R_i \cap S \neq \emptyset} R_i \supseteq S$$

And then we see that:

$$\begin{aligned}
U(\mathbb{1}_S, P) &= \sum_{i=1}^{m} M_{R_i}(\mathbb{1}_S) v(R_i) \\
&= \sum_{R_i \cap S \neq \emptyset} M_{R_i}(\mathbb{1}_S) v(R_i) \\
&= \sum_{R_i \cap S \neq \emptyset} v(R_i) = m(E_2) \\
L(\mathbb{1}_S, P) &= \sum_{i=1}^{m} m_{R_i}(\mathbb{1}_S) v(R_i) \\
&= \sum_{R_i \subseteq S} m_{R_i}(\mathbb{1}_S) v(R_i) \\
&= \sum_{R_i \subseteq S} v(R_i) = m(E_1)
\end{aligned}$$

Therefore!

$$m(E_2 \setminus E_1) = m(E_2) - m(E_1) = U(\mathbb{1}_S, P) - L(\mathbb{1}_S, P) < \varepsilon$$

Since $\varepsilon > 0$ was arbitrary, we conclude that S is Jordan measurable.

$2 \iff 3)$ This is straightforward using our characterization of integrability and the fact that $\mathbb{1}_S$ is discontinuous exactly at the points on the boundary of S.

$3 \implies 4)$ Let $\varepsilon > 0$. Since ∂S has Lebesgue measure zero there is a collection of boxes $B_1, B_2, \ldots$ such that $\partial S \subseteq \bigcup_{j=1}^{\infty} B_j$ and $\sum v(B_j) < \frac{\varepsilon}{2}$. Dilate each B_j into a larger open box $\widetilde{B}_j$ such that $B_j \subseteq \widetilde{B}_j$ and $v(\widetilde{B}_j) < 2v(B_j)$.

Now note that the $\widetilde{B}_j$ forms an open cover of the closed and bounded set ∂S.

By compactness there is a finite sub-cover $\widetilde{B}_{j_1}, \ldots, \widetilde{B}_{j_k}$ of ∂S. But then:

$$\sum_{i=1}^{k} v(\widetilde{B}_{j_k}) \leq \sum_{j=1}^{\infty} v(\widetilde{B}_j) < 2 \sum_{j=1}^{\infty} v(B_j) < \varepsilon$$

Great! This shows that ∂S has Jordan outer measure zero.

$4 \implies 3)$ follows trivially.

Improper Integrals

Up until now in the discussion of $\int_S f$ we restricted to the case where f and S are both bounded. In this section we relax these assumptions a bit to include any open set S and any continuous function f.

Remark. The ultimate dispensing of those two restrictions on S and f comes through the theory of Lebesgue integration.

Before we proceed, we introduce some notation:

- Let $\mathcal{J}$ denote the family of Jordan measurable subsets of $\mathbb{R}^n$.
- Let $\mathcal{J}_c$ denote the collection of compact Jordan measurable sets
- For a function $f : S \to \mathbb{R}$ we define the positive part and negative part of f as:

  $$f_+(x) = \max(f(x), 0) \qquad\qquad f_-(x) = \max(-f(x), 0)$$

  It is easy to veritfy that:

  - $f = f_+ - f_-$
  - $f_+, f_- \geq 0$
  - $|f| = f_+ + f_-$.
  - If f is continous then both f_+ and f_- are continuous.

Definition. *Let A be an open subset of $\mathbb{R}^n$ and let $f : A \to \mathbb{R}$ be a continuous function*

- *If f is non-negative on A we define the (extended) integral of f over A as:*

$$\int_A f = \sup_{\substack{\mathscr{D}\subseteq A \\ \mathscr{D}\in\mathcal{J}_c}} \int_{\mathscr{D}} f$$

  provided that this supremum exists.

- *If f is an arbitrary continuous function on A, write $f = f_+ - f_-$, where these are the positive and negative part of f. Provided that f_+ and f_- are integrable on A in the extended sense we say f is also integrable and let:*

$$\int_A f = \int_A f_+ - \int_A f_-$$

Remark. We now have two difference definitons of $\int_A f$ when A is open and bounded and f is continuous and bounded. We shall see later that these two definitions are equivalent if both integrals exist. The extended integral might exist without having the traditional integrals existince Why?

Notice that if $B \subseteq A$ are both open then if the extended integral of f over A exists then the extended integral of f over B exists and:

$$\int_B f \leq \int_A f$$

However if $f = 1$ then $\int_B 1$ exisrts only when B is Jordan measurable, and tehre are bounded open sets that are not Jordan measurable (we'll see an example in our Friday sessions)

Convention: If A is open and f is continuous then $\int_A f$ will always denote the extended integral

Lemma. *Let $A \subseteq \mathbb{R}^n$ be open. There exists a sequence of $C_1, C-2, \ldots$ of compact Jordan measurable sets such that $A = \bigcup_{i=1}^{\infty} C_j$ and $C_j \subseteq C_{j+1}^{\circ}$. In fact, C_j can be taken to be elementary*

Proof. Define:

$$\mathscr{D}_N = \{x \in \mathbb{R}^n \mid d(x, A^c) \geq \frac{1}{N}, |x| \leq N\}$$

Thus $\mathscr{D}_N$ is bounded and closed since $x \mapsto d(x, A^c)$ and $x \mapsto |x|$ are both continuous functions. Now consider:

$$A_{N+1} = \{x \in \mathbb{R}^n \mid d(x, A^c) > \frac{1}{N+1}, |x| < N+1\}$$

And then A_{N+1} is open and:

$$\mathscr{D}_N \subseteq A_{N+1} \subseteq \mathscr{D}_{N+1}$$

This implies that:

$$\mathscr{D}_N \subseteq \mathscr{D}_{N+1}^\circ$$

We clearly have by the fact that A is open that:

$$A = \bigcup_{N=1}^{\infty} \mathscr{D}_N$$

The sets $\mathscr{D}_N$ may not be Jordan measurable. To fix this, note that for $x \in \mathscr{D}_N$ there exists a closed cube centered at x and contained in $\mathscr{D}_{N+1}^\circ$. The interior of these cubes is an open cover of $\mathscr{D}_N$ and hence by compactness there is a fintie subcover. Define C_N to be the elementary set given by the finite union of such a finite subcover of $\mathscr{D}_N$ made up of closed cubes. Thus C_N is closed and bounded, and furthermore:

$$D_N \subseteq C_N^\circ \subseteq C_N \subseteq \mathscr{D}_{N+1}^\circ \subseteq C_{N+1}^\circ$$

Therefore we see that C_N is compact and Jordan measurable as well as the fact that $\bigcup_{N=1}^{\infty} C_N = A$. Great! This finishes the proof.

Theorem. *Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \to \mathbb{R}$ be a continuous function. Choose a sequence $C_N \in \mathcal{J}_c$ such that $A = \bigcup_{N=1}^{\infty} C_N$ and $C_N \subseteq C_{N+1}^\circ$ as in the above lemma. Then f is integrable over A if and only if $\int_{C_N} |f|$ is bounded by a constant which does not depend on N. In this case,*

$$\int_A f = \lim_{N \to \infty} \int_{C_N} f$$

In particular, f is integrable over A if and onl if $|f|$ is too.

We'll prove this theorem next time. In the meantime, here are some properties of the extended integral. For setup let $A \subseteq \mathbb{R}^n$ be open and let $f, g : A \to \mathbb{R}$ be continuous functions such that $\int_A f$ and $\int_A g$ exist:

a) $f + cg$ is integrable for any $c \in \mathbb{R}$ and:

$$\int_A f + cg = \int_A f + c \int_A g$$

b) If $f \leq g$ then:

$$\int_A f \leq \int_A g$$

In particular:

$$\left| \int_A f \right| \leq \int_A |f|$$

c) If A and B are both open and $A \subseteq B$ then if f is integrable over B then f is integrable over A. Furthermore if f is non-negative on B then:

$$\int_A f \leq \int_B f$$

d) If A and B are open and f is continuous on $A \cup B$, then if f is integrable on A and B then f is integrable on $A \cup B$ and $A \cap B$. Furthermore we have:

$$\int_{A \cup B} f = \int_A f + \int_B f - \int_{A \cap B} f$$

Handout 10

Where we are right now?

- **Lebesgue outer measure:** We modify the notion of Jordan outer measure by replacing the finite union of boxes by a countable union of boxes, i.e.
$$m^*(E) = \inf_{E\subset \cup_{j=1}^{\infty} B_j} \sum_{j=1}^{\infty} |B_j|$$
where the union above is taken over boxes $B_j \subset \mathbb{R}^d$. We saw last time that this is smaller than the Jordan outer measure and that the boxes above can be taken to be open or closed. We also saw that any countable set has zero Lebesgue outer measure.

- **Lebesgue measurability** A set $E \subset \mathbb{R}^d$ is said to be Lebesgue measurable if for every $\epsilon > 0$, there exists an open set $U \subset \mathbb{R}^d$ containing E such that $m^*(U \setminus E) \leq \epsilon$. If E is measurable, we refer to $m(E) = m^*(E)$ as the Lebesgue measure of E.

  We saw last time some properties of this definition:

  - Show that $m^*(\emptyset) = 0$.
  - (Monotonicity) Show that if $E \subset F \subset \mathbb{R}^d$, then $m^*(E) \leq m^*(F)$.
  - (Countable subadditivity) If $E_1, E_2, \ldots \subset \mathbb{R}^d$ is a countable sequence of sets, then $m^*\left(\cup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} m^*(E_n)$.

  A natural question is whether one has that an additivity property for the outer measure: namely that if E, F are <u>disjoint</u> sets then $m^*(E \cup F) = m^*(E) + m^*(F)$? While this turns out to be correct for some sets E and F (to be called Lebesgue-measurable sets), we already saw at the start of our discussion of measures that this cannot hold for general sets (cf. the Banach-Tarski paradox). The enemy here is that we might have the two sets E and F too intertwined or entangled together which can cause the additivity property to fail.

Q1) Show that if $\operatorname{dist}(E, F) > 0$, then $m^*(E \cup F) = m^*(E) + m^*(F)$.

Q2) Show that if E is an elementary set, then $m^*(E) = m(E)$ where $m(E)$ is the elementary measure of E defined before.

Q3) Conclude that if E is any bounded set, then $\underline{m}(E) \leq m^*(E) \leq \overline{m}(E)$ where $\underline{m}(E)$ and $\overline{m}(E)$ are the inner and outer Jordan measures of E.

Q4) Construct a bounded open subset U of $\mathbb{R}$ that is not Jordan measurable. *Hint: Start with an enumeration of the rationals in $[0, 1]$ and create an open set whose Lebesgue outer-measure is arbitrarily small but the Jordan outer measure is ≥ 1.*

MATH 395 Notes

Faye Jackson

November 20, 2020

Exercise 1. *Show that if* $\operatorname{dist}(E, F) > 0$ *then* $m^*(E \cup F) = m^*(E) + m^*(F)$*.*

Proof. We already have that $m^*(E \cup F) \leq m^*(E) + m^*(F)$. We now use the property of greatest lower bound to prove that $m^*(E \cup F) \geq m^*(E) + m^*(F)$. To do so, we will first prove a lemma:

Lemma. *For any sets* E *and* F *with* $\operatorname{dist}(E, F) > 0$ *and any box* B *we have that there is a finite collection of disjoint sub-boxes* $B_1, \ldots, B_N$ *covering* B *such that each* B_i *intersects at most one of* E *and* F*.*

Proof. Let $\varepsilon := \operatorname{dist}(E, F) > 0$. Now since $\varepsilon > 0$ we know that we can split B into sub-boxes $B_1, \ldots, B_N$ each of diameter less than ε. Then consider that for any i and any two points $X, y \in B_i$ we have:

$$d(x, y) \leq \operatorname{diam}(B_i) < \varepsilon = \operatorname{dist}(E, F)$$

We then may say that we cannot have $x \in E$ and $y \in F$, since if we did then we would have:

$$d(x, y) \leq \operatorname{diam}(B_i) < \varepsilon = \operatorname{dist}(E, F) \leq d(x, y)$$

Which is a contradiction. Therefore B_i intersects at most one of E and F.

Fix some countable collection $B_1, B_2, \ldots$ which covers $E \cup F$. We wish to show that $m^*(E) + m^*(F)$ is a lower bound for these, that is:

$$m^*(E) + m^*(F) \leq \sum_{i=1}^{\infty} B_i$$

Now for each B_i we use the lemma to split it into disjoint sub-boxes $B_{i1}, \dots, B_{iN_i}$ covering B such that each box B_{ij} intersects at most one of E and F. In particular we can split this up into disjoint collections of a countbale covering of E and a countable covering of F. Then by infimums:

$$\begin{aligned}
\sum_{i=1}^{\infty} B_i &= \sum_{i=1}^{\infty} \sum_{j=1}^{N_i} B_{ij} \\
&\geq \sum_{i=1}^{\infty} \sum_{\substack{j=1 \\ B_{ij} \cap E \neq \emptyset}}^{\infty} B_{ij} + \sum_{i=1}^{\infty} \sum_{\substack{j=1 \\ B_{ij} \cap F \neq \emptyset}}^{\infty} B_{ij} \\
&\geq m^*(E) + m^*(F)
\end{aligned}$$

Taking the infimum on the left hand side we see that:

$$m^*(E \cup F) \geq m^*(E) + m^*(F)$$

And therefore since we already have the other direction of the inequality by finite subadditivity we have $m^*(E \cup F) = m^*(E) + m^*(F)$ just as desired! Great!

Exercise 2. *Show that if E is an elementary set, then $m^*(E) = m(E)$ where $m(E)$ is the elementary measure of E defined before*

Proof. We want to only work with closed elementary sets. To do this we need a lemma:

Lemma. *For any elementary set E we have that $m^*(E)$ and $m^*(\overline{E})$.*

Proof. This is not too difficult. First note since $E \subseteq \overline{E}$ we have by monotonicity that $m^*(E) \leq m^*(\overline{E})$.

Now we wish to show that $m^*(E) \geq m^*(\overline{E})$. Note by finite sub-additivity we know:

$$m^*(\overline{E}) = m^*(E \cup \partial E) = m^*(E) + m^*(\partial E)$$

But wait! We know by previous IBL work that:

$$0 \leq m^*(\partial E) \leq \overline{m}_J(\partial E) = 0$$

Since we have previously shown that the Jordan measure of the boundary of a Jordan measurable set is zero, and E is elementary so it is Jordan measurable. But then

TODO

Now write $\overline{E}$, which must be an elementary set, as a finite union of disjoint boxes $E_1, \ldots, E_n$ by definition of an elementary set. Then note that the collection $E_1, \ldots, E_n$ covers $\overline{E}$, and so by definition of the Lebesgue outer measure as an infimum:

$$m^*(\overline{E}) \leq \sum_{j=1}^{n} |E_j| = m(\overline{E})$$

We now simply need to show the other inequality. To do so, it suffices to show that $m(\overline{E})$ is a lower bound for the set which defines $m^*(\overline{E})$ by the definition of infimum. By last homework, it suffices to consider countable coverings by open boxes.

Fix some countable collection of open boxes $B_1, B_2, \ldots$ which covers $\overline{E}$. Now consider that $\overline{E}$ is compact since elementary sets are bounded. Therefore there is a finite subcollection $B_1, \ldots, B_N$ which covers $\overline{E}$. By finite sub-additivity of the elementary measure:

$$m(\overline{E}) \leq \sum_{j=1}^{N} m(B_j) \leq \sum_{j=1}^{\infty} |B_j|$$

And therefore taking an infimum on the right hand side:

$$m(\overline{E}) \leq m^*(\overline{E})$$

But wait! Then by the lemma and previous work on elementary measure we have:

$$m^*(E) = m^*(\overline{E}) = m(\overline{E}) = m(E)$$

Great! This is exactly what we wanted to show!!! ☺

Exercise 3. ***TODO***

Proof. **TODO**

Exercise 4. ***TODO***

Proof. **TODO**

MATH 395 Notes

Faye Jackson

November 30, 2020

Announcements

- Final to be released on Monday December 14 in the afternoon, and due on Tuesday early morning. Say 4pm-4am
- To be submitted through gradescope

Recalling Improper Integrals

Recall. For A an open set and f continuous on A. We defined the extended $\int_A f$ as follows:

- If $f \geq 0$ then we define:

$$\int_A f = \sup_{\substack{D \in \mathcal{J}_c \\ D \subseteq A}} \int_D f$$

  Where $\mathcal{J}_c$ is the set of all compact Jordan measurable sets.

- For general f we write $f = f_+ - f_-$ and define:

$$\int_A f := \int_A f_+ - \int_A f_-$$

By convention if f is continuous and A is open then $\int_A f$ will mean the extended integral.

Problem: If A is open and bounded and f is continuous and bounded, we have two definitions for $\int_A f$. The extended integral may exist without having the ordinary integral existing. We will see today that if the ordinary integral exists then the extended integral exists and they are equal. We also proved the following

Lemma. *If $A \subseteq \mathbb{R}^n$ is open then there exists a sequence $C_1, C_2, \ldots$ of elementary sets (also compact Jordan measurable) such that:*

$$C_n \subseteq C_{n+1}^\circ$$
$$A = \bigcup_{j=1}^{\infty} C_j$$

Theorem. *Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \to \mathbb{R}$ be continuous. Choose a sequence $C_n \in \mathcal{J}_c$ as in the above lemma. Then f is integrable on A (in the extended sense) if and only if $\int_{C_n} |f|$ is bounded (uniformly in n). In this case,*

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

In particular, f is integrable on A if and only if $|f|$ is too.

Proof. We'll do this in cases:

- Let f be non-negative. In this case $\int_{C_n} f \, \mathrm{d}x$ is a monotonically increasing sequence of non-negative numbers, and as such it converges as $n \to \infty$ if and only if it is uniformly bounded.

  ($\Rightarrow$) Suppose that f is integrable over A. We want to show that $\int_{C_n} f$ exists and converges to $\int_A f$ as $n \to \infty$. Since f is continuous and C_n is compact, then f is bounded on C_n, and hence $\int_{C_n} f$ exists since C_n is Jordan measurable.

  Also:

  $$\int_{C_n} f \leq \sup_{\substack{\mathbb{D}\subseteq A \\ D\in\mathcal{J}_c}} \int_D f = \int_A f$$

  Therefore $\int_{C_n} f$ is uniformly bounded in n. This implies that it converges, now we need to show it converges to the right thing. We must also have that:

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

Great!

($\Leftarrow$) Suppose $\lim_{n\to\infty} \int_{C_n} f$ exists. Then $\int_{C_n} f$ is uniformly bounded in n by some constant M. Now take any $D \subseteq A$ and $D \in \mathcal{J}_c$. Then we know that:

$$D \subseteq \bigcup_{n=1}^{\infty} C_j^\circ$$

By compactness of D there exists a finite subcover, and since $C_j \subseteq C_{j+1}^\circ$ there exists some $n_\heartsuit$ such that $D \subseteq C_{n_\heartsuit}^\circ$. Therefore we know that:

$$\int_D f \leq \int_{C_{n_\heartsuit}} f \leq M$$

And therefore we have a nonempty bounded set, so the supremum exists:

$$\int_A f = \sup_{\substack{D \subseteq A \\ D \in \mathcal{J}_c}} \int_D f \leq M$$

Since M can be taken to be the limit as $n \to \infty$ of $\int_{C_n} f$ then we get that:

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

Combining these two inequalities from the if and only if we win and get the equality:

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

Perfect!

- Let's deal with general $f : A \to \mathbb{R}$ that is continuous. f is integrable over A if and only if f_+ and f_- are integrable if and only if $\int_{C_n} f_+$ and $\int_{C_n} f_-$ are bounded sequences by case one.

  But this is if and only if $\int_{C_n} f_+ + f_-$ is a bounded sequence, since $f_+, f_- \geq 0$. But since $f_+ + f_- = |f|$ this is only when $\int_{C_n} |f|$ is a bounded sequence. Therefore applying case 1 this is if and only if $\int_A |f|$ exists.

In this case we of course have:

$$\begin{aligned}
\int_{C_n} f_+ &\to \int_A f_+ \\
\int_{C_n} f_- &\to \int_A f_- \\
\int_{C_n} f &= \int_{C_n} f_+ - \int_{C_n} f_- \\
&\to \int_A f_+ - \int_A f_- \\
&= \int_A f
\end{aligned}$$

So we are done!

Theorem. *Let A be a bounded open set in $\mathbb{R}^n$ and let $f : A \to \mathbb{R}$ be a bounded continuous function. Then:*

a) The extended integral exists

b) If the ordinary integral exists, then the two integrals are equal.

Proof.

- Let us first show that the extended integral exists. Let M be an upper bound for $|f|$ on A. If $D \in \mathcal{J}_c$ is a subset of A, then:

  $$\int_D |f| \leq M \int_D 1 = Mv(D) \leq Mv(B)$$

  Where B is any box containing A. Therefore the set defining the extended integral is bounded, and so the extended integral of $|f|$ over A exists. This of course implies that the extended integral of f over A exists by our previous theorem.

- Now suppose that the ordinary integral $\int_A f$ exists and that $f \geq 0$. Then let B be a box containing A, then:

  $$(\text{ord}) \int_A f = \int_B f_A$$

Now let $D \subseteq A$ and $D \in \mathcal{J}_c$ then we must have that:

$$\int_D f = \int_D f_A \leq \int_B f_A = (\text{ord}) \int_A f$$

Therefore taking a sup over all D we get that:

$$(\text{ext}) \int_A f \leq (\text{ord}) \int_A f$$

To show the reverse inequality, let P be any partition of B and let $R_1, \ldots, R_m$ denote the sub-boxes of this partition. Now let $D = \bigcup_{R_i \subseteq A} R_i$. Then $D \subseteq A$ and $D \in \mathcal{J}_c$. Therefore:

$$\begin{aligned}
L(f_A, P) &= \sum_{i=1}^{m} m_{R_i}(f_A) v(R_i) \\
&= \sum_{R_i \subseteq A} m_{R_i}(f_A) v(R_i) \\
&\leq \sum_{R_i \subseteq A} \int_{R_i} f = \int_D f \\
&\leq (\text{ext}) \int_A f
\end{aligned}$$

Take the supremum over all such P and we obtain:

$$(\text{ord}) \int_A f = \sup_P L(f_A, P) \leq (\text{ext}) \int_A f$$

These two inequalities imply that the ordinary and extended integrals agree as desired to give (b) when $f \geq 0$.

- Write $f = f_+ - f_-$ as usual. Since f is integrable over A in the ordinary sense, so are $f_+ = \max(f, 0)$ and $f_- = \max(-f, 0)$. Therefore:

$$\begin{aligned}
(\text{ord}) \int_A f &= (\text{ord}) \int_A f_+ - (\text{ord}) \int_A f_- \\
&= (\text{ext}) \int_A f_+ - (\text{ext}) \int_A f_- \\
&= (\text{ext}) \int_A f
\end{aligned}$$

And this finishes the proof

Corrolary. *Let S be any bounded set and $f : S \to \mathbb{R}$ be a bounded continuous function. If f is integrable on S in the ordinary sense, then:*

$$\text{(ord)} \int_S f = \text{(ext)} \int_{S^\circ} f$$

Proof. Recall that if $\int_S f = \int_{S^\circ} f$, then apply the previous theorem.

This corollary is useful to translate results for extended integrals to ordinary integrals (like the change of variable formula in the next section).

The Change of Variables Formula

Recall. THe change of variable formula in 1D, otherwise known as u-substitution. Letting $f, g : [a, b] \to \mathbb{R}$ be functions with g C^1 and f continuous. Then letting $u = g(x)$ and $\mathrm{d}u = g'(x)\,\mathrm{d}x$ we have:

$$\int_a^b f(g(x))g'(x)\,\mathrm{d}x = \int_{g(a)}^{g(b)} f(u)\,\mathrm{d}u$$

There's a nuance, we are using $\int_a^b f$ to denote the <u>signed</u> integral which is defined as:

$$\int_a^b f = \begin{cases} \int_{[a,b]} f & \text{if } a \leq b \\ -\int_{[b,a]} f & \text{if } b < a \end{cases}$$

This u-substitution holds basically due to the chain rule, since if F is an antiderivative for f then $(F \circ g)' = f(g(x)) \cdot g'(x)$

Integrating from a to b then gives u-substitution by the Fundamental Theorem of Calculus.

There is no notion of signed integrals in higher dimensions, so we first need to formulate this theorem without signed integrals. For this note that $g([a, b]) = [g(a), g(b)]$ if g is increasing, i.e. $g' \geq 0$. And also $g([a, b]) = [g(b), g(a)]$ if g is decreasing, i.e. $g' \leq 0$.

If g is increasing then we can write:

$$\int_{[a,b]} f(g(x)) \cdot g'(x) = \int_{g([a,b])} f(u)\,\mathrm{d}u$$

And if g is decreasing then we can write:

$$\int_{[a,b]} f(g(x))g'(x)\,\mathrm{d}x = -\int_{g([a,b])} f(u)\,\mathrm{d}u$$

That is:

$$\int_{[a,b]} f(g(x))(-g'(x))\,\mathrm{d}x = \int_{g([a,b])} f(u)\,\mathrm{d}u$$

In either case, we may write that if g is monotone, then:

$$\int_{[a,b]} f(g(x))\left|g'(x)\right|\mathrm{d}x = \int_{g([a,b])} f(u)\,\mathrm{d}u$$

This is the formula that generalizes easily to higher dimensions.

So we look at this genralizing this via the correspondence:

1D	higher dimension
$[a,b]$	set A
$g([a,b])$	$g(A)$
g is monotone and C^1	g is a C^1 diffeomorphism
$u = g(x)$ $\mathrm{d}u = \lvert g'(x)\rvert\,\mathrm{d}x$	$u = g(x)$ $\mathrm{d}u = \lvert\det Dg\rvert\,\mathrm{d}x$

And so we have something like:

$$\int_A f(g(x))\,|\det Dg|\,\mathrm{d}x = \int_{g(A)} f(u)\,\mathrm{d}u$$

And we use this in the same way with:

$$u = g(x)$$
$$\mathrm{d}u = |\det Dg|\,\mathrm{d}x$$

Definition. *Let A be open in $\mathbb{R}^n$ and let $g : A \to \mathbb{R}^n$ be a one-to-one function of class C^r such that* $\det Dg(x) \neq 0$ *for $x \in A$. We call such a g a change of variables* *on A*

Remark. Recall that a C^r diffeomorphism is a one-to-one and onto function such that g and g^{-1} are in C^r

The inverse function theorem tells us that $g^{-1} \in C^r$ if $g \in C^r$ and $\det Dg(x) \neq 0$.

A change of variables on A is then nothing but a C^r diffeomorphism from A to $g(A)$

Theorem (Change of Variables Theorem)**.** *Let $g : A \to B$ be a C^1-diffeomorphism of open sets in $\mathbb{R}^n$ and let $f : B \to \mathbb{R}$ be a continuous function. Then f is integrable over B if and only if $f(g(x)) \cdot |\det Dg(x)|$ is integrable over A, and:*

$$\int_A f(g(x)) \cdot |\det Dg(x)| \, \mathrm{d}x = \int_B f(u) \, \mathrm{d}u$$

MATH 395 Notes

Faye Jackson

December 2, 2020

Change of Variables Theorem

Theorem. *We look at:*

$$\int_A f(g(x))\,|\det Dg(x)|\,\mathrm{d}x = \int_{g(A)} f(u)\,\mathrm{d}u$$

Intuitively we have:

$$\begin{aligned} u &= g(x) \\ \mathrm{d}u &= |\det Dg|\,\mathrm{d}x \\ x &\in A, u \in g(A) \end{aligned}$$

And so this holds whenever:

- $g : A \to g(A) = B$ *is a* C^1*-diffeomorphism*
- ***TODO***

Example. We look at Polar Coordinate Integration. Let:

$$B = \{(x, y) \in \mathbb{R}^2 \mid a^2 < x^2 + y^2 < b^2\}$$

Then there are the polar coordinates:

$$g(r, \theta) = (r\cos(\theta), r\sin(\theta))$$

Note that $B = g(A)$ where $A = \{(r, \theta) \mid a < r < b, 0 \leq \theta \leq 2\pi\}$. Then let us introduce:

$$\begin{aligned}\widetilde{A} &:= \{(r,\theta) \mid a < r < b, 0 < \theta < 2\pi\}\\ \widetilde{B} &:= g(\widetilde{A}) = B \setminus (\text{x-axis})\end{aligned}$$

And so then we have:

$$\begin{aligned}\int_{\widetilde{B}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_{g(\widetilde{A})} f(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_{\widetilde{A}} f(g(r,\theta)) \cdot |\det Dg(r,\theta)|\,\mathrm{d}r\,\mathrm{d}\theta\end{aligned}$$

And we know by previous homework that:

$$\begin{aligned}Dg(r,\theta) &= \begin{pmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{pmatrix}\\ \det Dg(r,\theta) &= r > 0\end{aligned}$$

Since we know that Dg is locally a C^1-diffeomorphism via the inverse function theorem and it is a bijection we know that it is a C^1-diffeomorphism, which is great. Now we apply Fubini:

$$\int_{\widetilde{B}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^{2\pi}\int_a^b f(r\cos\theta, r\sin\theta)r\,\mathrm{d}r\,\mathrm{d}\theta$$

Now since the x-axis has Lebesgue measure zero in $\mathbb{R}^2$, we then know that:

$$\int_B f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\widetilde{B}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^{2\pi}\int_a^b f(r\cos\theta, r\sin\theta)r\,\mathrm{d}r\,\mathrm{d}\theta$$

We know this because for C_N a nested sequence compact Jordan measurable set contained in B and covering B we know:

$$\begin{aligned}\int_{\widetilde{B}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \lim_{N\to\infty}\int_{C_N\setminus(\text{x-axis})} f(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \lim_{N\to\infty}\int_{C_N} f(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_B f(x,y)\,\mathrm{d}x\,\mathrm{d}y\end{aligned}$$

Great!

Example. Now for Spherical coordinate integration! Suppose we have:

$$B = \{(x, y, x) \mid x > 0, y > 0, z > 0, x^2 + y^2 + z^2 < a^2\}$$

Suppose we want to evaluate $\int_B f(x, y, z)\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z$. Suppose we take the change of coordinates:

$$\begin{aligned} x &= \rho \sin\phi \cos\theta \\ y &= \rho \sin\phi \sin\theta \\ z &= \rho \cos\phi \end{aligned}$$

And we'll denote this by $g(\rho, \phi, \theta)$. We aready calculated in previous homework that $\det Dg = \rho^2 \sin\phi$, and this is greater than 0 if $\rho > 0$ and $0 < \phi < \pi$. This happens on the set:

$$A = \left\{(\rho, \phi, \theta) : 0 < \rho < a, 0 < \phi < \frac{\pi}{2}, 0 < \theta < \frac{\pi}{2}\right\}$$

And here we have $g(A) = B$. Therefore using that g is a C^1 diffeomorphism from A to B and using Fubini we have that:

$$\begin{aligned} \int_B f(x, y, z)\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z &= \int_{g(A)} f(x, y, z)\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z \\ &= \int_A f(\rho \sin\phi \cos\theta, \rho \sin\phi \sin\theta, \rho \cos\phi)\rho^2 \sin\phi\, \mathrm{d}\rho\, \mathrm{d}\phi\, \mathrm{d}\theta \\ &= \int_0^a \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} f(\rho \sin\phi \cos\theta, \rho \sin\phi \sin\theta, \rho \cos\phi)\rho^2 \sin\phi\, \mathrm{d}\theta\, \mathrm{d}\phi\, \mathrm{d}\rho \end{aligned}$$

Some mapping Properties of diffeomorphisms

Lemma. *Let $A \subseteq \mathbb{R}^n$ be open and let $g : A \to \mathbb{R}^n$ be a C^1 function. If $E \subseteq A$ is a set of Lebesgue measure zero, then $g(E)$ also has Lebesgue measure zero.*

Remark. This is not true if g is only assumed to be continuous. In fact, there exists a continuous $g : [0, 1] \to [0, 1]^2$ that is onto. This is called Peano's space filling curve.

Proof. Let C_N be a family of compact sets such that $A = \bigcup_{N=1}^{\infty} C_N$ and $C_N \subseteq C_{N+1}^{\circ}$.

The note that:

$$E_N := E \cap C_N \qquad\qquad E = \bigcup_{N=1}^{\infty} E_N$$

It is enough to show that each $g(E_N)$ has Lebesgue measure zero.

Fix $\varepsilon > 0$ and let $M := \sup_{C_{N+1}} \|Dg\|_{\text{op}} < \infty$, since $g \in C^1$ and C_{N+1} is compact.

Also since $C_N \subseteq C_{N+1}^{\circ}$ there exists a $\delta > 0$ such that the δ-neighborhood of C_N is a subset of C_{N+1}°.

Since E_N has Lebesgue mreasure zero we can cover E_N by countably many boxes B_j such that $\sum v(B_j) < \varepsilon$. In fact, we can assume Without Loss of Generality that all the B_j are cubes and have diameter $< \delta$ by covering them with cubes of diameter $< \delta$.

Then $g(E_N)$ is a subset of $\bigcup g(B_j)$ where B_j is a cube of diameter less than δ

Claim. $\operatorname{diam} g(B_j) \leq M \operatorname{diam} B_j$.

Proof. Let $x, x' \in B_j$. By the Mean Value Theorem for some c on the line segment between x and x':

$$\begin{aligned} g(x) - g(x') &= Dg(c)(x - x') \\ \left|g(x) - g(x')\right| &\leq \|Dg(c)\|_{\text{op}} \left|x - x'\right| \\ &\leq M \operatorname{diam} B_j \end{aligned}$$

Great!

Therefore $g(B_j)$ is contained in a ball of radius $M \operatorname{diam} B_j$ which is then contained in a cube of $\widetilde{Q}_j$ of side length $2M \operatorname{diam} B_j$. Also:

$$\begin{aligned} \sum_j v(Q_j) &= \sum_j (2M)^n \cdot (\operatorname{diam} B_j)^n \\ &= \sum_j (2M)^n \cdot (v(B_j))^n \cdot C \end{aligned}$$

For some constant C, since the B_j are cubes, and so their diameter is proportional to their volume. But then:

$$\sum_j (Q_j) = C(2M)^n \cdot \sum_j v(B_j) < (2M)^n \cdot C \cdot \varepsilon$$

But then since $(2M)^n \cdot C$ is a constant, we can take $\varepsilon \to 0$ and we will be done. This finishes the proof.

Corrolary. *Let $g : A \to B$ be a diffeomorphism betweeen two open sets A and B. Let $K \subseteq A$ be compact. Then:*

a) $g(K^\circ) = (g(K))^\circ$ and $g(\partial K) = \partial g(K)$

b) If K is Jordan measurable, then so is $g(K)$.

These results hold if K is not compact provided that $\partial K \subseteq A$ and $\partial g(K) \subseteq B$.

Proof. Let's go!

a) This takes some work!

- Since g^{-1} is continuous, then g is open. Therefore if $B(x, \delta) \subseteq K$ then $g(B(x, \delta))$ is an open subset of $g(K)$, which implies that $g(B(x, \delta)) \subseteq (g(K))^\circ$. And so $g(K^\circ) \subseteq (g(K))^\circ$.
- Also $g(A \setminus K) \subseteq B \setminus g(K)$ since g is one-to-one. Let $y \in \partial g(K)$. Then there exists an $x \in A$ such that $y = g(x)$. We know that $x \notin K^\circ$ since then y would belong to $(g(K))^\circ$.

  We also know $x \notin A \setminus K$ since otherwise $y \in B \setminus g(K)$ which also does not intersect $\partial g(K)$ since $g(K)$ is closed. Therefore $x \in \partial K$, and so $\partial g(K) \subseteq g(\partial K)$.
- Apply the same argument to g^{-1} and $g(K)$ to obtain that:

$$\begin{aligned} g^{-1}((g(K))^\circ) &\subseteq K^\circ \\ \partial K &\subseteq g^{-1}(\partial g(K)) \end{aligned}$$

  And therefore:

$$\begin{aligned} (g(K))^\circ &\subseteq g(K^\circ) \\ g(\partial K) &\subseteq \partial g(K) \end{aligned}$$

Combining this with the previous part gives part (a)

b) Note that if K is Jordan measurable, then ∂K has Lebesgue measure zero. Since g is C^1 we then know that $g(\partial K) = \partial g(K)$ has Lebesgue measure zero, and so $g(K)$ is Jordan measurable.

Volumes and Determinants

Theorem. *Let A be an $n \times n$ matrix and let $h : \mathbb{R}^n \to \mathbb{R}^n$ be the transformation $h(x) = Ax$. Let S be a Jordan measurable set in $\mathbb{R}^n$ and $T := h(S)$. Then:*

$$v(T) = |\det A|\, v(S)$$

Proof. T is Jordan measurable by the above corollary. Therefore when $|\det A| \neq 0$ we have by the change of variables that:

$$\begin{aligned} v(T) = v(T^\circ) &= \int_{T^\circ} 1 \,\mathrm{d}x \\ &= \int_{h(S^\circ)} = \int_{S^\circ} |\det A| \,\mathrm{d}y \\ &= |\det A|\, v(S^\circ) = |\det A|\, v(S) \end{aligned}$$

In Case 2, when $\det A = 0$ we know that the range of h is a subspace V of $\mathbb{R}^n$ os dimension $p < n$. Since V has Lebesgue measure zero (check!), we are done, since then $T \subseteq V$ will have Lebesgue measure zero.

MATH 395 Notes

Faye Jackson

December 7, 2020

Last time, we proved that:

Theorem. *Let A be an $n \times n$ matrix and $h : \mathbb{R}^n \to \mathbb{R}^n$ be given by $h(x) = A \cdot x$. If S is Jordan measurable then $h(S)$ is Jordan measurable and:*

$$\mathrm{vol}(h(S)) = |\det A| \cdot \mathrm{vol}(S)$$

Corrolary. *Let $a_1, \dots, a_n$ be n linearly independent vectors of $\mathbb{R}^n$. Let $A = [a_1, \dots, a_n]$ be the $n \times n$ matrix whose columns are $a_1, \dots, a_n$ and let P be the parallelopiped given by:*

$$P = \left\{ \sum c_i a_i \mid 0 \le c_i \le 1 \right\}$$

Then $v(P) = |\det A|$

Proof. Let $h(x) = Ax$, then h takes the unit cube in $\mathbb{R}^n$ to P. Therefore:

$$\mathrm{vol}(P) = \mathrm{vol}(h(S)) = |\det A| \cdot \mathrm{vol}([0,1]^n) = |\det A|$$

Orientations

Definition. *Let $\beta = (a_1, \dots, a_n)$ be a basis of $\mathbb{R}^n$. We call this basis right-handed if $\det(a_1, \dots, a_n) > 0$ and left-handed if $\det(a_1, \dots, a_n) < 0$.*

On a general vector space V. Let $\beta = (v_1, \dots, v_n)$ and $\beta' = (w_1, \dots, w_n)$ be two bases of V. Let $w_j = a_{j1} v_1 + \dots + a_{jn} v_n$. Then the matrix $A = (a_{jk})$ is invertible *since:*

$$A = {}_{\beta'}[\mathrm{Id}]_{\beta}$$

is a change of basis matrix. We say that β and β' have the same orientation if $\det A > 0$ *and opposite orientation if* $\det A < 0$.

Remark. The choice of notation is motivated by the 2D and 3D cases in which we have the right-hand rule

Exercise. *Show that:*

1) *This gives an equivalence relation on the set of bases of V with two equivalence classes.*

2) *Another way to define this equivalence relation is as follows. Pick $T : \mathbb{R}^n \to V$ a linear isomorphism. Any basis β of V can be writte as $\{Ta_1, \ldots, Ta_n\}$ where $(a_1, \ldots, a_n)$ is a basis of $\mathbb{R}^n$. So given two bases $\beta = \{Ta_1, \ldots, Ta_n\}$ and $\beta' = \{Tb_1, \ldots, Tb_n\}$.*

   β and β' have the same orientation if and only if $(a_1, \ldots, a_n$ and $(b_1, \ldots, b_n)$ have the same orientation in $\mathbb{R}^n$.

Theorem. *Let C be a non-singular $n \times n$ matrix and let $h : \mathbb{R}^n \to \mathbb{R}^n$ be given by $h(x) = Cx$. Let $(a_1, \ldots, a_n)$ be a basis in $\mathbb{R}^n$. Then the two bases $(a_1, \ldots, a_n)$ and $(h(a_1), \ldots, h(a_n))$ have the same orientation if and only if* $\det C > 0$.

Proof. Let $b_j = h(a_j)$. Then $C[a_1, \ldots, a_n] = [b_1, \ldots, b_n]$. But then:

$$\det C \cdot \det(a_1, \ldots, a_n) = \det(b_1, \ldots, b_n)$$

And so $\det C > 0$ if and only if $\det(a_1, \ldots, a_n)$ and $\det(b_1, \ldots, b_n)$ have the same sign, which is exactly when they have the same orientation.

Isometries of $\mathbb{R}^n$

Definition. *Let $h : X \to Y$ be a map between metric spaces (X, d_X) and (Y, d_Y). We say that h is an isometry provided that:*

$$d_Y(h(x_1), h(x_2)) = d_X(x_1, x_2) \qquad (x_1, x_2 \in X)$$

Remark. Isometries are always one-to-one, but they might not be onto. For example $h : \mathbb{R} \to \mathbb{R}^2$ where $h(x) = (x, 0)$.

Here we will discuss isometries from $\mathbb{R}^n \to \mathbb{R}^n$ with the same Euclidean metrix

Example. Lets grab some examples!

1) Consider $h : \mathbb{R}^n \to \mathbb{R}^n$ where $h(x) = x - a$ for a constant $a \in \mathbb{R}^n$, since:

$$h(x) - h(y) = x - a - y + a = x - y \implies \|h(x) - h(y)\| = \|x - y\|$$

2) Let $h : \mathbb{R}^n \to \mathbb{R}^n$ where $h(x) = Ax$ and A is an orthogonal matrix. Then h is an isometry:

   Recall. A is orthogonal means $A^T A = AA^T = \mathrm{Id}$. In other words:

$$\langle Ax, Ay\rangle = \langle A^T Ax, y\rangle = \langle x, y\rangle$$

   That is A preserves inner products

   But then we know that:

$$\begin{aligned} \|Ax - Ay\|^2 &= \langle Ax - Ay, Ax - Ay\rangle \\ &= \langle A(x - y), A(x - y)\rangle \\ &= \langle x - y, x - y\rangle = \|x - y\|^2 \end{aligned}$$

   And therefore h is an isometry.

The interesting fact is that these are the only two examples of isometries on $\mathbb{R}^n$

Theorem. *Let $h : \mathbb{R}^n \to \mathbb{R}^n$ be a map such that $h(0) = 0$. Then:*

a) h is an isometry if and only if h preserves inner products

b) h is an isometry if and only if $h = Ax$ where A is an orthogonal matrix.

Proof. Let's go!

a) Consider that:

$$\begin{aligned} \|h(x) - h(y)\|^2 &= \langle h(x) - h(y), h(x) - h(y)\rangle \\ &= \langle h(x), h(x)\rangle - 2\langle h(x), h(y)\rangle + \langle h(y), h(y)\rangle \end{aligned}$$

And:

$$\begin{aligned}\|x-y\|^2 &= \langle x-y, x-y\rangle \\ &= \langle x,x\rangle - 2\langle x,y\rangle + \langle y,y\rangle\end{aligned}$$

Now we can do this. Therefore if h preserves inner products we must have $\|h(x)-h(y)\| = \|x-y\|$, and so h is an isometry.

On the other hand, if h is an isometry and $h(0)=0$ then:

$$\langle h(x), h(x)\rangle = |h(x)|^2 = |h(x)-h(0)|^2 = |x-0|^2 = \langle x,x\rangle$$

We also know for every $x, y \in \mathbb{R}^n$ that $|h(x)-h(y)|^2 = |x-y|^2$ and so using the above two equations again we see that:

$$2\langle h(x), h(y)\rangle = 2\langle x,y\rangle \implies \langle h(x), h(y)\rangle = \langle x,y\rangle$$

b) The backwards implication was discussed in the previous direction. For the forward direction consider $\{h(e_1), h(e_2), \ldots, h(e_n)\}$ where $e_1, \ldots, e_n$ is the standard basis of $\mathbb{R}^n$. Since h preserves inner products $\{h(e_1), \ldots, h(e_n)\}$ is an orthonormal set, which implies that it is an orthonormal basis.

Therefore for any $x \in \mathbb{R}^n$ we can express:

$$h(x) = \sum_{j=1}^{n} \alpha_j(x) h(e_j)$$

And then we know that:

$$\begin{aligned}\langle h(x), h(e_k)\rangle &= \left\langle \sum_{j=1}^{n} \alpha_j(x) h(e_j), h(e_k)\right\rangle \\ &= \sum_{j=1}^{n} \alpha_j(x) \cdot \langle h(e_j), h(e_k)\rangle \\ &= \sum_{j=1}^{n} \alpha_j(x) \cdot \langle e_j, e_k\rangle \\ &= \alpha_k(x)\end{aligned}$$

But then we have that:

$$\alpha_k(x) = \langle h(x), h(e_k) \rangle = \langle x, e_k \rangle = x_k$$

And therefore:

$$h(x) = \sum_{j=1}^{n} x_j h(e_j) = Ax$$

where $A = [h(e_1), \ldots, h(e_n)]$. Since this is an orthonormal basis, A is orthogonal and so we are done.

Corrolary. *Let $h : \mathbb{R}^n \to \mathbb{R}^n$. Then:*

1) h is an isometry if and only if it is an orthogonal transformation followed by a translation. I.e. $h(x) = Ax + p$ where A is an orthogonal matrix and $p \in \mathbb{R}^n$.

2) If h is an isometry, then h preserves volumes as well. That is if S is Jordan measurable, then $h(S)$ is Jordan measurable and:

$$v(h(S)) = v(S)$$

Proof. This is pretty cool!

1) Let $\widetilde{h}(x) = h(x) - h(0)$. Then h is an isometry if and only if $\widetilde{h}$ is an isometry with $\widetilde{h}(0) = 0$, and this holds by the previous theorem if and only if $\widetilde{h}(x) = Ax$ for A some orthogonal matrix.

   Then by rearrangement h is an isometry if and only if:

$$h(x) = \widetilde{h}(x) + h(0) = Ax + h(0)$$

   For some orthogonal matrix A.

2) We know that $A \cdot S$ is Jordan measurable with volume $|\det A| \cdot v(S) = v(S)$ since $|\det A| = 1$ when A is orthogonal. Of course $A \cdot S + p$ has the same measure as $A \cdot S$, and so $h(S) = A \cdot S + p$, and therefore $v(h(S)) = v(S)$ as desired!!!

Great!