Analysis I

Thomas Fleming

November 4, 2021

Contents

Lecture 1: Overview and Basics

Tue 24 Aug 2021 12:50

1 Overview and Basics

Definition 1.1 (Functions)**.** A function $f : X \to Y$ consists of a **domain**, X, a **codomain** Y, and the **range**, $f(X) = \{f(x) \in Y : x \in X\}$. The image of a set $A \subseteq X$ is $f(A) = \{f(x) \in Y : x \in A\}$. The **preimage** of a set $B \subseteq Y$ is the set $f^{-1}(B) = \{x \in X : f(x) \in B\}$. The **restriction** of f to $C \subseteq X$ is $f|_C : C \to Y$ such that $f|_C(x) = f(x)$ for $x \in C$. The **composition** of f with $g : Y \to Z$ is $(g \circ f)(x) = g(f(x))$.

Definition 1.2 (Properties of Functions)**.** A function $f : X \to Y$ is called **surjective** or **onto** if $f(X) = Y$. f is **one-to-one** or **injective** if $f(x) = f(y) \Rightarrow x = y$ for $x, y \in X$. f is **bijective** if it is one-to-one and onto.

Example. The function $id_X : X \to X : x \mapsto x$ is called the **identity**.
If $f : X \to Y$ is injective, then there is a function $g : f(X) \to X$ such that $(g \circ f) = id_X$. We call g the **inverse** of X and denote it f^{-1}.

$\diamond$

Definition 1.3 (Sequences)**.** Given $m \in \mathbb{N}$, a function $f : \{k \in \mathbb{N} : 1 \leq k \leq m\} \to X$ is called a **finite sequence** of length m. This is typically denoted by $f(k) = (x_k)_{k=1}^{m} = (x_k)_{1 \leq k \leq m} = *(x_1, x_2, \ldots, x_m)$.
A function $f : \mathbb{N} \to X$ is called a **sequence** or **infinite sequence**. This is typically denoted $f(k) = (x_k) = (x_k)_k = (x_k)_{k=1}^{\infty} = (x_k)_{1 \leq k}$ based on context.
A sequence need not have first index 1 be it finite or infinite, in this case we denote it, for first index m and last index n, $f(k) = (x_k)_{k=m}^{k=n} = (f_k)_{m \leq k \leq n}$.
Similarly for infinite sequences of first index m, we denote $f(k) = (x_k)_{m \leq k}$.

Definition 1.4 (Collections)**.** A **family** or **collection** of subsets of X is a subset of $\mathscr{P}(X) = \{A : A \subseteq X\}$.
A collection $\mathscr{C}'$ is a **subcollection** of the collection $\mathscr{C}$ if $\mathscr{C}' \subseteq \mathscr{C}$.

Definition 1.5 (Indexed Collection)**.** Let $f : \Lambda \to \mathscr{P}(X)$. f is called an **indexed collection** of subsets of X for the indexed set Λ of subsets of X. We generally denote this by $f(\lambda) = (A_\lambda : \lambda \in \Lambda) = (A_\lambda)$ based on context.

Definition 1.6 (Set Operations)**.** We denote the **intersection** $A \cap B = \{x : x \in A \text{ and } x \in B\}$,
the **union** $A \cup B = \{x : x \in A \text{ or } x \in B\}$,
the **complement** $A^C = X \setminus A = \{x : x \notin A \text{ and } x \in X\}$,
the **symmetric difference** $A \Delta B = (A \setminus B) \cup (B \setminus A)$.
We say sets A, B are **disjoint** if $A \cap B = \varnothing$.
A collection $\mathscr{C}$ of subsets of X is a **disjoint collection** if for all $A, B \in \mathscr{C}$, A and B are disjoint unless $A = B$.
Let $\mathscr{C}$ be a collection of subsets of X, we define $\bigcap_{A \in \mathscr{C}} A = \{x \in X : x_n A \ \forall \ A \in \mathscr{C}\}$.
Similarly, we define $\bigcup_{A \in \mathscr{C}} A = \{x \in X | \exists \ A \in \mathscr{C} \text{ such that } x \in A\}$.
Lastly, the special case $\mathscr{C} = \{A_k : 1 \leq k \leq m\}$ or $\{A_k : k \in \mathbb{N}\}$ allows the notation $\bigcap_{k=1}^{m} A_k$ or $\bigcap_{k=1}^{\infty} A_k$ and $\bigcup_{k=1}^{m} A_k$ or $\bigcup_{k=1}^{\infty} A_j$

Definition 1.7 (De Morgen's Laws)**.** Let $\mathscr{C} \subseteq \mathscr{P}(X)$. First we state the **negation laws**, $\left(\bigcup_{A\in\mathscr{C}} A\right)^C = \bigcap_{A\in\mathscr{C}} A^C$ and $\left(\bigcap_{A\in\mathscr{C}} A\right)^C = \bigcup_{A\in\mathscr{C}} A^C$ Now, the **distributive laws**, let $B \subseteq X$, then $B \cap \left(\bigcup_{A\in\mathscr{C}} A\right) = \bigcup_{A\in\mathscr{C}} (A \cap B)$ and $f^{-1}\left(\bigcap_{B\in\mathscr{C}'} B\right) = \bigcap_{B\in\mathscr{C}'} f^{-1}(B)$.

Proposition 1.1. Let $\mathscr{C} \subseteq \mathscr{P}(X)$, $\mathscr{C}' \subseteq \mathscr{P}(Y)$, $A \subseteq X$, $B \subseteq Y$, $f : X \to Y$.
Then, $f\left(\bigcup_{A\in\mathscr{C}} A\right) = \bigcup_{A\in\mathscr{C}} f(A)$. However, $f\left(\bigcap_{A\in\mathscr{C}} A\right) \neq \bigcap_{A\in\mathscr{C}} f(A)$.
Also, $f^{-1} * \left(\bigcup_{B\in\mathscr{C}'} B\right) = \bigcup_{B\in\mathscr{C}'}$ and $f^{-1}\left(\bigcap_{B\in\mathscr{C}'} B\right) = \bigcap_{B\in\mathscr{C}'} f^{-1}(B)$.
Next, $f^{-1}\left(B^C\right) = \left[f^{-1}(B)\right]^C$.
Last, $f\left(f^{-1}(B)\right) \subseteq B$ and $A \subseteq f^{-1}(f(A))$.

Definition 1.8 (Countability)**.** X is **finite** if X is empty of if there is $m \in \mathbb{N}$ and a bijection function $f : X \to \{k | 1 \leq k \leq m\}$. If X is not finite, it is called **infinite**.
X is called **countably infinite** if there is a bijection $g : X \to \mathbb{N}$. X is called **countable** if it is finite or countably infinite.
X is called **uncountable** if it is not countable.

Example. $\mathbb{N}, \mathbb{Z}$, and $\mathbb{Q}$ are countable, but $\mathbb{R}$ is uncountable.

◇

Proposition 1.2. 1. Every subset of a countable collection is countable.
2. The set of all finite sequences from a countable set is countable.
3. The union of a countable collection of countable sets is countable.

Definition 1.9 (Algebras)**.** A collection $\mathscr{A}$ of subsets of X is called an **algebra** if all of the following hold:
1. $X \in \mathscr{A}$,
2. $A \in \mathscr{A} \Rightarrow A^C \in \mathscr{A}$,
3. $A \cup B \in \mathscr{A}$ for all $A, B \in \mathscr{A}$.

Definition 1.10 (σ-Algebra)**.** A collection $\mathscr{A}$ of subsets of X is called a **σ-algebra** if all of the following hold:
1. $X \in \mathscr{A}$,
2. $A \in \mathscr{A} \Rightarrow A^C \in \mathscr{A}$,
3. $\bigcup_{k=1}^{\infty} A_k \in \mathscr{A}$ for every (countable) collection of subsets $\{A_k : k \in \mathbb{N}\}$.

Lecture 2: Algebras

Thu 26 Aug 2021 13:02

2 Algebras and Measure Theory

For an interval (a, b) we define the length of the interval to be $b - a$. Similarly, for an interval extending to infinity in either direction, we define the length to be infinite. Again, for the union of multiple intervals, we simply define their length by the sum of the respective interval lengths. We wish to generalize this notion, and for this reason we develop measure theory.

Remark. There are sets which are not measurable. The construction of such sets will come later.

We will begin defining which sets are measurable by constructing σ-algebras. The simplest σ-algebra is formed by taking the power set of a given collection.

Proposition 2.1. For every collection $\mathscr{C}$ of subsets of X, there is a smallest algebra $\mathscr{A} \supseteq \mathscr{C}$. That is, if $\mathscr{D}$ is an algebra such that $\mathscr{D} \supseteq \mathscr{C}$, then $\mathscr{D} \supseteq \mathscr{A}$. Note this applies to σ-algebras as well. In the case of σ-algebra, we call this the **borrel σ-algebra**.

Proof. Let $\mathscr{F}$ be the set of all algebras (σ-algebras) which contain $\mathscr{C}$. Note, as $\mathscr{P}(X)$ is always an algebra, then we see $\mathscr{P}(X) \in \mathscr{F}$, hence $\mathscr{F}$ is nonempty. Now, let $\mathscr{A} = \bigcap_{B \in \mathscr{F}} B$. Then, as $\mathscr{C} \subseteq B$ for all $B \in \mathscr{F}$, we have $\mathscr{C} \subseteq \mathscr{A}$. Similarly, $X \in \mathscr{A}$, because $X \in \mathscr{B}$ for all $B \in \mathscr{F}$. If $A \in \mathscr{A}$, then $A \in B$ for all $B \in \mathscr{F}$. Since every B is an algebra (σ-algebra), then $A^c \in B$ for all $B \in \mathscr{F}$, hence $A^c \in \bigcap_{B \in \mathscr{F}} B = \mathscr{A}$. If $\{A_k\}$ is a finite (countable) collection with $A_k \in \mathscr{A}$, then $A_k \in B$ for all $B \in \mathscr{F}$. Since each B is an algebra (σ-algebra), then $\bigcup_k A_k \in B$ for all $B \in \mathscr{F}$, then $\bigcup_k A_k \in \bigcap_{B \in \mathscr{F}} B = \mathscr{A}$. It is clear this is the smallest algebra as it is the smallest set containing every element of $\mathscr{F}$. $\square$

Remark. σ-algebras are the domains of measures. That is, if given a measure, it must be defined over a σ-algebra.

Definition 2.1 (Choice Function)**.** Given a collection $\mathscr{F}$ of nonempty sets, a **choice function** $f : \mathscr{F} \to \bigcup_{F \in \mathscr{F}} F$, $F \mapsto f(F) \in F$. For example, let $\mathscr{F} = \{\{1, 2, 3\}, \{2, 5\}, \{1\}\}$, then $f(\{1, 2, 3\}) = 1, 2,$ or 3.

Of course, such a choice function is simple to construct for a finite number of countable sets, but consider $\mathscr{F} = \mathscr{P}(\mathscr{P}(\mathscr{P}(\mathbb{R})))$, then it is very unclear how one would go about selecting an element from each set.

Law 1 (Axiom of Choice)**.** For all collections $\mathscr{F}$, there is always a choice function (which may be defined non-constructively).

While this axiom is not necessary to construct measure theory, its exclusion makes certain basic theorems unprovable, for instance the proposition from last lecture that the countable union of countable sets is countable. Generally, we work with a weaker version of the axiom of choice:

Law 2 (Zermelo's Axiom of Choice)**.** Every collection of nonempty sets has a choice function.

Other weakenings include the axiom of countable choice. The name Zermelo comes from Zermelo-Franklin set theory, known as ZF theory. When Zermelo's AOC is included in the axioms, we denote this model **ZFC set theory**.

Proposition 2.2. Let $\{X_\lambda\}$ be a collection of nonempty sets indexed by $\lambda \in \Lambda$. Let X be the cartesian product

$$\prod_{\lambda \in \Lambda} X_\lambda = \{g : g : \Lambda \to \bigcup_{\lambda \in \Lambda} X_\lambda : g(\lambda) \in X_\lambda \text{ for every } \lambda \in \Lambda\}.$$

Properly defining this product over arbitrary Λ requires some form of Axiom of Choice.

Remark. In the case of such a cartesian product, taking AOC, we know there is a choice function $f : \{X_\lambda : \lambda \in \Lambda\} \to \bigcup_{\lambda \in \Lambda} X_\lambda, X_\lambda \ \mapsto f(X_\lambda) \in X_\lambda$. Letting $g(\lambda) = f(X_\lambda)$ yields the desired g for $\prod_{\lambda \in \Lambda} X_\lambda$.

Definition 2.2 (Relation)**.** Let X be nonempty, then a **relation**, R, is a subset of $X \times X$. If an element $x \in X$ is in a relation with $y \in Y$ we could write $(x, y) \in R$, but this is nonstandard. The normal notation is xRy to mean x is in relation with y.

Definition 2.3 (Properties of a Relation)**.**

1. Reflexive: xRx for all $x \in X$.
2. Symmetric: $xRy \Rightarrow yRx$.
3. Transitive: If xRy and yRz, then xRz.

Example. The simplest example is equality,=, which is reflexive, symmetric, and transitive.
Another example is the order relation of $\mathbb{Q}$, $\leq$. This is reflexive, not symmetric, and transitive. $\diamond$

Definition 2.4 (Equivalence Relation)**.** An **equivalence relation** is a relation, R, which is reflexive, symmetric, and transitive. We generally denote an equivalence relation (other than equality) by $\sim$.

Remark. Because of the desired properties of equivalence relations, they allow us to partition a set X into **equivalence classes** or **components**. That is, given two equivalence classes $[x_1] = \{y \in X : y \sim x_1\} = [\hat{x_1}]$ if and only if $x_1 \sim \hat{x_1}$. Here x_1 or $\hat{x_1}$ or any other member of their equivalence class are called the representative. One last property is that given two representatives x, y either $[x] = [y]$ or $[x] \cap [y] = \varnothing$.

Definition 2.5 (Partial Order)**.**

1. A reflexive transitive relation R on a nonempty set X is called a **partial ordering** if xRy and yRx imply $x = y$.
2. A partial ordering R on a nonempty set X is called a **total ordering** if for all $x, y \in X$ we have at least one of xRy or yRx being true. In this case we call X a **total ordered** or sometimes just **ordered**.
3. For a nonempty X with partial ordering R we call $z \in X$ an **upper bound** of a subset $A \subseteq X$, for all $x \in A$, xRz is true.
4. For a nonempty X and a partial ordering R, we call $z \in X$ a **maximal element** if zRx implies $z = x$.

Example. $\leq$ is a complete order on $\mathbb{R}$, but when we extend $\leq$ to $\mathbb{C}$ such that $x \leq y$ is only well defined for $x, y \in \mathbb{R}$, it becomes only a partial ordering. $\diamond$

A classical example of a statement about orderings is Zorn's Lemma, which turns out to be equivalent to the Axiom of Choice.

Theorem 2.1 (Zorn's Lemma)**.** Let X be a nonempty set with a partial ordering. If every totally ordered subset of X has an upper bound, then X has a maximal element.

A very similar result to this is Hausdorff's Maximality Theorem.

Lecture 3: Construction of the Reals

Tue 31 Aug 2021 12:17

3 Construction of the Reals

Definition 3.1 (Rational Cauchy Sequences)**.** We define a **rational cauchy sequence** to be $(x_n)_{n\in\mathbb{N}}$ with $x_n \in \mathbb{Q}$, such that for all $\varepsilon > 0$, there is a $N \in \mathbb{N}$ such that $|a_n - a_m| < \varepsilon$ for $n, m \geq N$. We denote the set of all rational cauchy sequences to be $\mathrm{CS}(\mathbb{Q})$.

Just as with standard cauchy sequences, we have that $(a_k), (b_k) \in \mathrm{CS}(\mathbb{Q})$ implies $(a_k \pm b_k), (a_k \cdot b_k) \in \mathrm{CS}(Q)$.
Now, we define an equivalence relation on $\mathrm{CS}(\mathbb{Q})$ such that $(a_k) \sim (b_k)$ if for all $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that $|a_m - b_m| <$ if $m \geq N$.

Remark. It is of note that this essentially identifies all sequences which want to converge to the same point (though they may not converge to a rational number, or a number at all).

Now, we examine the equivalence classes, that being the quotient space $\mathrm{CS}(\mathbb{Q}) / \sim$, and we note that if $q \in \mathbb{Q}$, then $(q)_k \in [(q)_k]$. That is, we can identify $\mathbb{Q}$ with a subset of the quotient space.
Next, we aim to extend the total ordering of $\mathbb{Q}$ to $\mathrm{CS}(\mathbb{Q}) / \sim$. Let us introduce the relation $\leqslant$ on $\mathrm{CS}(Q)$ such that for $x, y \in \mathrm{CS}(Q) / \sim$, we say $x \leqslant y$ if for all $(x_k) \in x$ and $(y_k) \in y$ and for every rational $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $x_m \leq y_m + \varepsilon$ for $m \geq N$.

Proposition 3.1. Suppose $x, y \in \mathrm{CS}(\mathbb{Q}) / \sim$ and there are $(x_k) \in x$, $(y_k) \in y$ with the property that for all rational $\varepsilon > 0$ there is a $N \in \mathbb{N}$ such that $x_m \leq y_m + \varepsilon$ if $m \geq N$, then $x \leqslant y$. Essentialy, we need not show the property holds for all members of a given equivalence class, but merely a representative of a given equivalence class.

Proof. Let $\varepsilon > 0$ be rational, then there is $N \in \mathbb{N}$ such that $x_m \leq y_m + \frac{\varepsilon}{3}$ if $m \geq N$. For $(\hat{x_k}) \in x$ and $(\hat{y_k}) \in y$, then there is $K \in \mathbb{N}$ such that $|\hat{x_m} - x_m| < \frac{\varepsilon}{3}$ and $|\hat{y_m} - y_m| < \frac{\varepsilon}{3}$ if $m \geq K$.
Hence, if $m \geq \max\{K, N\}$, then

$$\hat{x_m} < x_m + \frac{\varepsilon}{3} \leq y_m < \frac{2\varepsilon}{3} < \hat{y_m} + \varepsilon.$$

Hence $\hat{x_m} < \hat{y_m} + \varepsilon$. □

Lemma 3.1. The relation $\leqslant$ is a total ordering on $\mathrm{CS}(\mathbb{Q}) / \sim$.

Proof. First, let us show it is a partial ordering. It is clear that $\leqslant$ is reflexive and transitive as it identifies equivalence classes. Now, suppose $x, y \in \mathrm{CS}(\mathbb{Q}) / \sim$ such that $x \leqslant y$ and $y \leqslant x$. Given $(x_k) \in x$ and $(y_k) \in y$, for a given rational $\varepsilon > -$ we can find $N \in \mathbb{N}$ such that $x_m \leq y_m + \frac{\varepsilon}{2}$ and $y_m \leq x_m + \frac{\varepsilon}{2}$ for $m \geq N$. Hence, $|x_m - y_m| \leq \frac{\varepsilon}{2} < \varepsilon$ for $m \geq N$, hence

$$\underbrace{[(x_m)]}_{=x} = \underbrace{[(y_m)]}_{=y}.$$

Thus, $x = y$, so $\leqslant$ is a partial ordering.

Now, let us show any two elements are comparably to obtain a total ordering. Let $x, y \in \mathrm{CS}(\mathbb{Q}) / \sim$ and suppose $x \nleqslant y$, then we must show $y \leqslant x$. For $(x_k) \in x$, $y_k \in y$ we have a rational $\varepsilon > 0$ such that $x_m > y_m + \varepsilon$ for infinitely many $m \in \mathbb{N}$. Also, there is $K \in N$ such that for $k, m \geq k$, we have $|x_k - x_m| < \frac{\varepsilon}{2}$ and $|y_k - y_m| < \frac{\varepsilon}{2}$. Hence, if $m \geq K$ such that $x_m > y_m + \varepsilon$ (there are infinitely many so we must only chooses a sufficiently large one), we obtain for $k \geq K$,

$$\begin{aligned} x_k + \frac{\varepsilon}{2} &> x_k + (x_m - x_k) \\ &= x_m \\ &> y_m + \varepsilon \\ &= y_k + (y_m - y_k) + \varepsilon \\ &> y_k + \varepsilon - \frac{\varepsilon}{2} \\ &= y_k + \frac{\varepsilon}{2}. \end{aligned}$$

Hence, $x_k > y_k$, so $y_k < x_k + \varepsilon$ for $\varepsilon > 0$ and $k \geq K$. Hence $y \leqslant x$, so $\leqslant$ is a total ordering. □

Definition 3.2 (Operations on $\mathrm{CS}(\mathbb{Q})/\sim$)**.** Now, we wish to introduce addition of $\mathrm{CS}(\mathbb{Q}) / \sim$.
Let $x, y \in \mathrm{CS}(\mathbb{Q}) / \sim$ with $(x_k) \in x$ and $(y_k) \in y$, we define $x \pm y = [(x_k \pm y_k)]$ and $xy = [(x_k y_k)]$.

We note that this is well defined as we already know $(x_k \pm y_k)$ and $(x_k y_k)$ to be valid rational cauchy sequences. Furthermore, these definitions are independent of which representative we choose. The proof of this is trivial by the algebraic properties of $\mathbb{Q}$.

Definition 3.3 (Real Numbers)**.** We define the **real numbers**, $\mathbb{R}$ to be $\mathrm{CS}(\mathbb{Q})$ together with the total ordering $\leqslant$, addition, subtraction, and multiplication.

If x is rational, integral, or natural, we simply write $x = x \in \mathbb{R}$ because we can identify x with its equivalence class. Furthermore, if $p, q \in \mathbb{Q}$ such that $p \leq q$, then $p \leqslant q$, hence we may say for $a, b \in \mathbb{R}$ with $a \leqslant b$ we also have $a \leq b$, so the notation $\leqslant$ becomes redundant. Furthermore, the other order symbols $<, >, \geq$ are also valid on $\mathbb{R}$, defined in their usual way.

Lemma 3.2 (The Archimidean Property)**.** For any two real numbers $x, y \in \mathbb{R}$ with $x > 0$, there exists a natural number m such that $mx > y$.

Proof. Suppose the opposite, then $mx \leq y$ for all $m \in \mathbb{N}$. Let $(x_k) \in x$ and $(y_k) \in y$ and let $M \in \mathbb{Q}$ such that $y_m \leq M$ for all $m \in \mathbb{N}$ (as rational cauchy sequences are always bounded). Since $\mathbb{Q}$ is archimidean (trivially), for given rational $\varepsilon > 0$, we find there exists $m \in \mathbb{N}$ such that $\frac{M+1}{m} < \varepsilon$. Also, there is $N \in \mathbb{N}$ such that $0 \leq m + \frac{\varepsilon}{2}$ for $m \geq N$ and $mx_m \leq y_m + 1$ for $m \geq N$.

Hence, the latter inequality implies $x_m \leq \frac{y_m+1}{m} \leq \frac{M+1}{m} < \varepsilon$ while the first inequality shows $x \geq -\frac{\varepsilon}{2} > -\varepsilon$ for all $m \geq N$. Consequently, $|x_m| < \varepsilon$ for $m \geq N$. Hence $[(x_m)] = x = 0$. $\lightning$. So, the claim is true. $\square$

Corollary 1. For any two real $x, y \in \mathbb{R}$ with $x < y$ there exists a rational $q \in \mathbb{Q}$ such that $x < q < y$.

Lecture 4: Construction of the Reals (2) and Intro to Topology

Thu 02 Sep 2021 13:05

Recall (Archimedian property)**.** For $x, y \in \mathbb{R}$ with $x > 0$, there is $m \in \mathbb{N}$ such that $mx > y$.

Corollary 2. For any two $x, y \in \mathbb{R}$ with $x < y$ there is a rational q such that $x < q < y$.

Proof. By Archmedian property there is a $m \in \mathbb{N}$ such that $m(y-x) > 1$. Additionally, there is $j \in \mathbb{N}$ such that $j > -mx$. Now, let $\ell \in \mathbb{N}$ be the smallest natural such that $\ell > mx + j$ (by the fact that rational cauchy sequences are always bounded, such an ℓ must exist and well-ordering guarantees a smallest such ℓ). Then,

$$\begin{aligned} my + j &> mx + 1 + j, \text{ by construction of } m \\ &\geq \ell, \text{ by construction of } \ell \\ &> mx + j. \end{aligned}$$

Hence, $my > \ell - j > mx$. Now, let $(x_k) \in x$, $(y_k \in y)$ be representatives of x and y respectively. Then, for every rational $\varepsilon > 0$ there is a $N \in \mathbb{N}$ such that $\ell - j \leq my_m + m\varepsilon$ and $mx_m \leq l - j + m\varepsilon$ for $m \geq N$. Hence, with $q = \frac{\ell - j}{m}$ we get $q \leq ym + \varepsilon$ and $x_m \leq q + \varepsilon$. This implies that $q \leq y$ and $x \leq q$.
Now, suppose $q = x$ or $q = y$, then $mx = \ell - j$ or $my = \ell - j$, but as $my > \ell - j > mx$, we see this cannot happen. Hence $x < q < y$. $\square$

Definition 3.4 (Inverses)**.** For $x \in \mathbb{R}$, we call $y \in \mathbb{R}$ the **inverse** of x if $xy = 1$.

Proposition 3.2. Every real $x \neq 0$ has a unique inverse.

Proof. For $x \neq 0$, let $(x_k) \in x$ be a representative. We wish to construct a cauchy sequence $\left(\frac{1}{x_k}\right)$ such that multiplication provides the constant (1) sequence. Of course, some x_k could be 0, but we know we can always choose a sufficiently small rational $0 < \varepsilon < x$ which allows us to guarantee for $n < N \in \mathbb{N}$, $x_k \neq 0$.
By the corollary, we know there is a $q \in \mathbb{Q}$ such that either $x < -2q < 0$ or $0 < 2q < x$. Hence there exists $N \in \mathbb{N}$ such that $2q \leq |x_n| + q$ if $n > N$ (i.e. $|x_n| \geq q$

for $q > 0$). Now, define (y_k) by setting $y_m = \begin{cases} \frac{1}{x_n}, & n \geq N \\ 1, & uad1 \leq n < N \end{cases}$ $1)Thisis1indeedarationalcauchysequen$
$(x_k y_k) = 1$. Hence (y_k) is an inverse. Now, we prove uniqueness.
Let $(z_k) \in \mathrm{CS}(\mathbb{Q})$ such that $[(x_k z_k)] = 1$. Then, $(x_k z_k) \sim 1 \sim (x_k y_k)$. Since $|x_k| \geq q$ for $k \geq N$, then we see

$$\begin{aligned} |y_m - z_m| &= |x_m y_m - x_m z_m| \cdot \frac{1}{|x_m|} \\ &\leq \frac{1}{q} |x_m y_m - x_m z_m| \text{ if } m \geq N. \end{aligned}$$

Consequently $[(y_k)] = [(z_k)]$. So, the inverses are unique. □

Notation. For $x \neq 0$ we denote the inverse of x to be x^{-1}.

Consequently, $\mathbb{R}$ is an ordered field. Hence, we may define division:

Definition 3.5 (Division)**.** We define, for $x, y \in \mathbb{R}$, with $y \neq 0$, $\frac{x}{y} = xy^{-1}$.

Now, we have finally finished constructing $\mathbb{R}$ as we have already known it, but we are missing the motivation for the usefulness of $\mathbb{R}$, as all of its properties so far are mirrored in $\mathbb{Q}$. This next property is unique to $\mathbb{R}$ and it is what makes it of unique interest.

Definition 3.6 (Bounds)**.**

1. $M \in \mathbb{R}$ is an **upper bound** of $S \subseteq R$ if $x \leq m$ for all $x \in S$.
2. $m \in \mathbb{R}$ is an **lower bound** of $S \subseteq R$ if $x \geq m$ for all $x \in S$.
3. $M^{\circledast} \in \mathbb{R}$ is the **least upper bound** or **supremum** of $S \subseteq R$ if $M^{\circledast}$ is an upper bound of S and $M^* \leq M$ for all other upper bounds M.
4. $m^{\circledast} \in \mathbb{R}$ is the **greatest lower bound** or **infimum** of $S \subseteq R$ if $M^{\circledast}$ is a lower bound of S and $M^* \geq M$ for all other lower bounds M.

We generally denote $\mathrm{LUB}(S) = \sup(S)$ and $\mathrm{GLB}(S) = \inf(S)$ for the supremum and infimum respectively.

Corollary 3. If S has no lower bound, we write $\sup(S) = \infty$, and if S has no upper bound, we write $\inf(S) = -\infty$. Additionally, $\sup(\varnothing) = -\infty$ and $\inf(S) = \infty$. Lastly, these bounds are unique.

Proposition 3.3. Let $S \subseteq \mathbb{R}$. An upper bound M^* is the least upper bound of S if and only if for every $\varepsilon > 0$ there is a $s \in S$ such that $M^* < s + \varepsilon$.

Proof. Suppose M^* is an upper bound such that for every $\varepsilon > 0$ there is a $s \in S$ such that $M^* > s + \varepsilon$. Let M be any with $M < M^*$ and let $\varepsilon = M^* - M$.Then, there is $s \in S$ such that $M^* < s + \varepsilon = s + M^* - M$, in other words $M < s$.

Hence, M is not an upper bound of S, so $M^* = \sup(S)$. Conversely, if M^* is the least upper bound, then for every $\varepsilon > 0$, $M^* - \varepsilon$ is not an upper bound of S. Hence, there is $s \in S$ such that $M^* - \varepsilon < s$ and hence $M^* < s + \varepsilon$. □

Of course, this also holds for lower bounds.

Theorem 3.1 (Least Upper Bound Property)**.** If a nonempty subset of $\mathbb{R}$ has an upper bound, it has a supremum.

This is the property which makes $\mathbb{R}$ of interest, and its proof will provide a generalizable format for proving other statements. We will halve intervals to construct a rational cauchy sequence such that its equivalence class will be the supremum.

Proof. Let $S \subseteq R$ be nonempty with an upper bound. As $\mathbb{R}$ is archimedian, there exist rational numbers ℓ_1, u_1 such that u_1 is an upper bound and l_1 is not (For every $s \in S$ we can find $\ell_1 < s < u_1$). Now, construct sequences $(\ell_m), (u_m)$of rationals such that each u_m is an upper bound and each ℓ_n is not. Having constructed ℓ_m, u_m we define ℓ_{m+1}, u_{m+1} such that $u_{m+1} = \frac{\ell_m + u_m}{2}$ and $\ell_{m+1} = \ell_m$ if this new u_{m+1} is an upper bound, and if not we set $u_{m+1} = u_m$ and $\ell_{m+1} = \frac{\ell_m + u_m}{2}$ as this will guarantee $\ell_q \leq s \leq u_q$ for all $s \in S$. Additionally, it follows that

$$\ell_1 \leq \ell_m \leq \ell_{m+1} \leq \ldots \leq u_{m+1} \leq u_m \leq u_1.$$

Since we have $u_{m+1} - \ell_{m+1} = \frac{1}{2}(u_m - \ell_m)$, we get for every rational $\varepsilon > 0$ and $k \geq m$

$$\begin{aligned} \max\{|\ell_k - \ell_m|, |u_k - u_m|\} &\leq u_m - \ell_m \\ &\leq |u_m - \ell_m| \\ &\leq \frac{2}{2^m}|u_1 - \ell_1|, \text{ if and only if } m \text{ is sufficiently large.} \end{aligned}$$

This shows that $(\ell_m), (u_m) \in \mathrm{CS}(\mathbb{Q})$. Let $\ell = [(\ell_m)]$ and $u = [(u_m)]$. Then, by the same inequality, we see $u = \ell$.
Next, we will show u is indeed an upper bound of S. Suppose the contrary, then there is $s \in S$ such that $u < s$. Let $\varepsilon = s - u > 0$. By construction of ℓ and u, we have $\ell_k \leq \ell = u \leq u_k$ and $|\ell_k - u_k| \leq \frac{2}{2^k}|\ell_1 - u_1|$ for every k. Hence, we fine $K \in \mathbb{N}$ such that $u_k - u \leq u_k - \ell_k < \varepsilon$. Thus $u_k < u + \varepsilon = u + s - u = u$. $\lightning$. So, we have that u is in fact an upper bound.
Finally, we show it is a least upper bound. Since for every $\varepsilon > 0$ there is $K \in \mathbb{N}$ such that $u - \ell_k \leq u_k - \ell_k < \varepsilon$, we have that $\ell_k \geq u_k - \varepsilon \geq u - \varepsilon$. Since ℓ_k is not an upper bound by construction, there is $s \in S$ such that $s > \ell_k \geq u_k - \varepsilon$ or $u < s + \varepsilon$. Hence, there can be no upper bound larger than u, so u is a least upper bound of S. □

The same line of reasoning yields a Greatest Lower Bound Property.

Remark. Any ordered field which has the LUB property yields a bijection between $\mathbb{R}$ and this field which preserves all structure. This is an important result as, up to structure preserving bijections, $\mathbb{R}$ is the unique ordered field with the Least Upper Bound Property.

4 Topology on $\mathbb{R}$

Definition 4.1 (Open/Closed)**.** A set $S \subseteq \mathbb{R}$ is **open** if for each $x \in S$ there is an $\varepsilon > 0$ such that $(x-\varepsilon, x+\varepsilon) \subseteq S$. A set $V \subseteq \mathbb{R}$ is **closed** if V^c is open.

Definition 4.2. Let $U, V \subseteq \mathbb{R}$ then,

1. $x \in \mathbb{R}$ is a **interior point** of U if there is $\varepsilon > 0$ such that $(x-\varepsilon, x+\varepsilon) \subseteq U$. The **interior** of a set $U \subseteq \mathbb{R}$, denoted $U^\circ = \overset{\circ}{U}$ is the set of all interior points of U.
2. $x \in R$ is a **closure point** of V if for every $\varepsilon > 0$, $V \cap (x-\varepsilon, x+\varepsilon) \neq \varnothing$. The **closure** of V is denoted $\overline{V} = \mathrm{cl}\,(V)$ is the set of all closure points of V.

Lecture 5: Topology on $\mathbb{R}$ (2)

Tue 07 Sep 2021 13:05

Recall. $U \subseteq \mathbb{R}$ is open iff $U = \overset{\circ}{U}$ and $V \subseteq \mathbb{R}$ is closed iff $V = \overline{V}$ where $\overline{V}$ is the closure of V.

Proposition 4.1.

1. $\mathbb{R}$ and $\varnothing$ are open.
2. $\bigcup_{x \in X} U_x$ is open for arbitrary open U_x and ordered x.
3. $\bigcap_{i=1}^{n} U_i$ is open for finite n an open U_i
4. The complements of these statements hold for closed sets.

Proposition 4.2. Let $U \subseteq \mathbb{R}$. Then $\overset{\circ}{U}$ is the largest ope set contained in U and $\overline{U}$ is the smallest closed set containing U.

Definition 4.3 (Boundary)**.** For any $S \subseteq \mathbb{R}$, the set $\delta S = \overline{S} \setminus \overset{\circ}{S}$. A point $x \in \delta S$ is called a **boundary point**.

Example. For $S = \mathbb{Q} \cap [0,1]$ we have $\overline{S} = [0,1]$ and $\overset{\circ}{S} = \varnothing$. Hence $\delta S = S$. ⋄

Lemma 4.1. For any set $S \subseteq \mathbb{R}$, $\left(\overline{S}\right)^c = (\overset{\circ}{S^c})$.

Theorem 4.1. Every open set in $\mathbb{R}$ is the disjoint union of a countable collection of open intervals.

Proof. Let $U \subseteq \mathbb{R}$ be open and denote $x \sim y$ if $[x, y] \subseteq U$ if $x \leq y$ or $[y, x] \subseteq U$

if $y \leq x$. It is trivial to show that this is infact an equivalence relation on U. Now analyzing the equivalence classes guarantees disjointness. Now, we must only show that the equivalence classes are intervals and are countable.
Denote the equivalence class containing x by I_x. As $x \in I_x$, we see each $I_x \neq \varnothing$. Furthermore, $I_x \subseteq U$ for all $x \in U$ and for any $x, \hat{x} \in U$ we have either $I_x = I_{\hat{x}}$ or $I_x \cap I_{\hat{x}} = \varnothing$.
Next, note that each I_x, $x \in U$ is an interval, else we could find a $y \in I_x$ and a $z \notin I_x$ such that one of the following occure

1. $x < z < y$ and $[x, y] \subseteq U$ but $[x, z] \not\subseteq U$ $\lightning$.
2. $y < z < x$ and $[y, x] \subseteq U$ but $[x, z] \not\subseteq U$ $\lightning$.

Hence, I_x is an interval. Next, suppose $y \in I_x$ for some $x \in U$. We wish to show $y \in \overset{\circ}{U}$. Since U is open there is an $\varepsilon > 0$ such that $(y - \varepsilon, y + \varepsilon) \subseteq U$. Cosequently, for each $z \in (y - \varepsilon, y + \varepsilon)$ we conclude $z \sim y \sim x$. Hence, $(y - \varepsilon, y + \varepsilon) \subseteq I_x$. Hence I_x is open for each $x \in U$.
Lastly, we show these unions are countable. This will invoke the axiom of choice so it is not valid in all models.
Note, $U = \bigcup_{x \in U} I_x$. Also, note that $I_x \cap \mathbb{Q} \neq \varnothing$ as each I_x is open and nonempty, hence convexity implies it contains a rational. Hence, there is a choice function $f : \{I_x \cap \mathbb{Q} : x \in U\} \to \bigcup_{x \in U} (I_x \cap \mathbb{Q}) = U \cap \mathbb{Q}$ such that $f(I_x \cap \mathbb{Q}) \in I_x \cap \mathbb{Q}$ for every $x \in U$.
Now, the map $g : \{I_x : x \in U\} \to R$ with R being the image of f and $g(I_x) = f(I_x \cap \mathbb{Q})$. This is trivially a bijection from $U \to R$, and as $R \subseteq \mathbb{Q}$ we have countability.
Essentially this argument first strips all the irrationals from I_x and then chooses 1 rational contained in I_x. The use of two functions here just cleans up the arguments, but essentially, f does all the work while g just adds the formalism to create a true bijection. $\square$

Remark. Picking a rational from each set and construcing a bijection $\alpha : U \to \mathbb{Q}$ is a common strategy.

Definition 4.4 (Compactness)**.** Let $S, U_\lambda \subseteq \mathbb{R}$ with $\lambda \in \Lambda$. Then, $\{U_\lambda : \lambda \in \Lambda\}$ is a **cover** of S if $S \subseteq \bigcup_{\lambda \in \Lambda} U_\lambda$. The cover is an **open cover** if each U_λ is open and it is a countable cover if Λ is countable.
A **subcover** of a cover $\{U_\lambda : \lambda \in \Lambda\}$ is a collection $\{U_\lambda : \lambda \in \Lambda_0\}$ such that $\Lambda_0 \subseteq \Lambda$ and $\{U_\lambda : \lambda \in \Lambda_0\}$ is a cover.
A set is called **compact** if every open cover of the set has a finite subcover.
A set S is **connected** if for all open sets U, V such that $S \subseteq U \cup V$ and $S \cap U \cap V = \varnothing$ it follows that $S \subseteq U$ or $S \subseteq V$.

Theorem 4.2 (Heine-Borel Theorem)**.** A set $S \subseteq \mathbb{R}$ is compact if and only if S is bounded and closed.

Proof. Suppose S is bounded and closed. Let us examine the special case, $S = [a, b] \subseteq \mathbb{R}$.
Let $\{U_\lambda : \lambda \in \Lambda\}$ be an open cover of S and let F be the set of all $x \in S$, with

$a \leq x \leq b$ such that $[a, x]$ has a finite subcover from $\{U_\lambda\}$. Clearly, $a \in F$ as $[a, a] = \{a\} \in U_\lambda$ for some λ. Since F is bounded from above b there exists a supremum of F. Let $c = \sup(F)$. Furthermore, $a \leq c \leq b$ by definition of supremum , so $c \in [a, b]$. Then, $c \in U_{\lambda_0}$ for some $\lambda_0 \in \Lambda$ as c is in the interval which is covered by $\{U_\lambda\}$. Since U_{λ_0} is open by assumption, there is a $\varepsilon > 0$ such that $(c - \varepsilon, c + \varepsilon) \subseteq U_x$. Since $c - \varepsilon$ is not an upper bound of F, there is an $x \in F$ such that $c - \varepsilon < x$. Since $x \in F$, there is a finite indexed set $\Lambda_0 \subseteq \Lambda$ such that $[a, x] \subseteq \bigcup_{\lambda \in \Lambda_0} U_\lambda$. Hence, $[a, c + \varepsilon] \subseteq \bigcup_{\lambda \in \Lambda_0 \cup \{U_{\lambda_0}\}} U_\lambda$. As Λ_0 is countable, $\Lambda_0 \cup \{U_{\lambda_0}\}$ yields a countable subcover. Hence, $c = b$, else c would not be an upper bound, so $[a, b] \subseteq [a, b + \varepsilon)$ has a finite subcover, so it is compact. Lastly, let S be any bounded, closed set, in $\mathbb{R}$. Then, there is $M \in \mathbb{N}$ such that $|x| \leq M$ for all $x \in S$, so $S \subseteq [-M, M]$ and we see an open cover of S, $\{U_\lambda : \lambda \in \Lambda\}$. Then, $\{U_\lambda : \lambda \in \Lambda\} \cup S^c = \mathbb{R}$ is an open cover of $[a, b]$, hence there is a finite subcover of $[a, b]$. If the finite subcover does not contain S^c, we have a finite subcover of S, otherwise, we can take the finite subcover minus S^c to induce a finite subcover of S.
Conversely, suppose S is not bounded, then $S \subseteq \bigcup_{k \in \mathbb{Z}} (k, k + 2)$ and we see any finite subcover of this would yield a bounded set. Similairly, if S is not closed, then there is a closure point $x \notin S$. Then, $S \subseteq \bigcup_{\varepsilon > 0} ([x - \varepsilon, x + \varepsilon])^c = \mathbb{R} \setminus \{x\}$. As all intervals $(x - e, x + e) \cap S \neq \varnothing$, we see any finite collection of ε will yield a minimum ε, denoted ε_0, hence a finite subcover would be of the form $\bigcup_{i=1}^n (x - \varepsilon_i, x + \varepsilon_i) = \mathbb{R} \setminus (x - \varepsilon_0, x + \varepsilon_0) \not\supseteq S$. $\lightning$. Hence, S is bounded and closed. $\square$

Corollary 4. Suppose $\{C_\lambda : \lambda \in \Lambda\}$ is a collection of compact sets such that $\bigcap_{\lambda \in \Lambda_0} C_\lambda \neq \varnothing$ for every finite indexed $\Lambda_0 \subseteq \Lambda$. Then $\bigcap_{\lambda \in \Lambda} C_\lambda \neq \varnothing$.

Proof. Assume $\bigcap_{\lambda \in \Lambda} C_\lambda = \varnothing$. Then, $\bigcup_{\lambda \in \Lambda} C_\lambda^c = \mathbb{R}$. Given $\Lambda_0 \in \Lambda$, we find a finite subcover $\{C_\lambda^c : \lambda \in \Lambda_0\}$ of C_{λ_0}. That is, $C_{\lambda_0} \subseteq \bigcup_{\lambda \in \Lambda_0} C_\lambda^c$ fora finite subset $\Lambda_0 \subseteq \Lambda$. Hence, $C_{\lambda_0} \bigcap_{\lambda \in \Lambda_0} C_\lambda = \varnothing$. $\lightning$. Hence, $\bigcap_{\lambda \in \Lambda} C_\lambda \neq \varnothing$. $\square$

Lecture 6: Sequences in $\mathbb{R}$

Thu 09 Sep 2021 13:03

Lemma 4.2. A set $S \subseteq \mathbb{R}$ is connected if and only if it is an interval.

Remark. $\varnothing = (a, a)$ is considered an interval as is a singleton $([x, x])$.

Proof. By the previous remark we may assume $|S| \geq 2$.
Now, assume S is connected and not an interval, then we can contain $x, y \in S$ such that $x < z < y$. Then, we can construct neighborhoods $U = (-\infty, z)$ and $V = (z, \infty)$ which are open and disjoint with $S \subseteq U \cup V$ and $S \cap V \cap U$ hence S is not connected.
Now, assume S is an interval. Let U, V be open such that $S \subseteq U \cup V$ and $S \cap U \cap V = \varnothing$. Suppose $x, y \in S$ with $x < y$. Furthermore, WLOG we may let $x \in U$. Now, let $\{z : [x, z] \subseteq U, x \leq z \leq y\}$ and as y is an upper bound of this set, we may let c be its supremum.
Then, $[x, c] \subseteq S$. Then $[x, c] \subseteq S$. Hence, $c \in U$ for if $c \in V$ and $(c - \varepsilon, c + \varepsilon) \subseteq V$ for some ε, then we would conclude that $S \cap U \cap v \neq \varnothing$. Since $c \in U$, then we

see $c = y$. Otherwise, we could find $\varepsilon > 0$ and $z \in U$ such that $c < z < c + \varepsilon \leq y$ and $[x, z] \in U$. Consequently $S \subseteq U$, so an interval is connected.
In short we showed that for every $x \in U$ and all $x < y \in U$ such that the $[x, y] \subseteq U$ hence U contains all intervals in S, hence $S \subseteq U$. $\square$

Definition 4.5 (Borel σ-Algebra)**.** The smallest σ-Algebra containing all open subsets of $\mathbb{R}$ is called the **Borel σ-Algebra**. We denote F_σ to be the union of a countable collection of closed sets and G_δ to be the intersection of a countable collection of open sets.

Definition 4.6.

1. A sequence (x_m) is **bounded** if there is a $M > 0$ such that $|x_m| \leq M$ for all m.
2. A sequence (x_m) is **bounded from above or below** if there is an $M \in \mathbb{R}$ such that $x_m \leq M$ or $x_m \geq M$ for all m.
3. A sequence is **increasing** if $x_n \leq x_{n+1}$ for all n, it is **strictly increasing** if $x_n < x_{n+1}$.
4. A sequence is **decreasing** if $x_n \geq x_{n+1}$ and similarly for **strictly decreasing** we say $x_n > x_{n+1}$
5. A sequence (x_m) is called **convergent** if there is $x_0 \in \mathbb{R}$ such that for every $\varepsilon > 0$ we find that $N \in \mathbb{N}$,
$$|x_m - x_0| < \varepsilon \text{ if } m \geq N.$$
This is writen $\lim x_n, \lim_n x_n,$ or $\lim_{n\to\infty} x_n$.
6. A sequence (x_n) has a **subsequence** $(x_m)_k$. All sequences have a strictly increasing and a strictly decreasing subsequence.
7. For a sequence (x_m) and a point $x \in \mathbb{R}$, x is an **accumulation/cluster point** of (x_m) if for every $\varepsilon > 0$ the set $\{m : |x_m - x| < \varepsilon\}$ is infinite. Lastly, if a limit does converge, its limit is unique.

Example. For the sequence $x_m = (-1)^m$ we see $1, -1$ are accumulation points.
The sequence $x_m = (1 + e^{-m})$ yields accumulation point 1. $\diamond$

Proposition 4.3. A point x^* is an accumulation point of the sequence (x_n) if and only if (x_m) has a convergent subsequence if and only if (x_m) has a convergent subsequence with limit x^*.

Proposition 4.4. Let $S \subseteq \mathbb{R}$ be nonempty. Then, a point $x^* \in \mathbb{R}$ belongs to $\overline{S}$ if and only if there is a convergent sequence, $(x_n) \in S^{\mathbb{N}}$ with $x^* = \lim_{n\to\infty} x_n$.

Theorem 4.3 (Bolzano-Weirstrass Theorem)**.** Every bounded sequence has a convergent subsequence.

Proof. Given a bounded sequence (x_n), let $M > 0$ be a bound. We will construct intervals $[A_k, B_k]$ such that the set $\{n : x_n \in [A_k, B_k]\}$ is infinite and $[A_{k+1}, B_{k+1}] \subseteq [A_k, B_k]$ and $B_k - A_k \leq \frac{4M}{2^k}$.
Set $A_1 = -M$, $B_1 = M$. Having constructed A_k, B_k, define A_{k+1}, B_{k+1} as follows

- If $\{m : x_m \in \left[A_k, \frac{A_k+B_k}{2}\right]\}$ is infinite, let $A_{k+1} = A_k$ and $B_{k+1} = \frac{A_k+B_k}{2}$.
- Otherwise, let $A_{k+1} = \frac{A_k+B_k}{2}$ and $B_{k+1} = B_k$.

It is clear the intervals are nested and infinite, as we always took either the top or bottom half of an infinite interval, one of which must be infinite. Since $[A_{k+1}, B_{k+1}] \subseteq [A_k, B_k]$ for all k, we can always take a finite number of sets and find a finite intersection. Then, by the finite intersection property, $\bigcap_{k=1}^{\infty} [A_k, B_k] \neq \varnothing$.
Suppose $x_0, y_0 \in \bigcap_{k=1}^{\infty} [A_k, B_k]$ then for every $k \in \mathbb{N}$ and $\varepsilon > 0$, $|x_0 - y_0| \leq B_k - A_k \leq \frac{4M}{2^k < \varepsilon}$ for sufficiently large k. Hence $x_0 = y_0$.
Lastly, we construct a subsequence which converges to x_0 . Let $n_1 = 1$ and having found n_k, note that the construction guarantees the set $\{m : x_m \in [A_{k+1}, B_{k+1}], m > n_k\}$ is an infinite set. By well-ordering, it contains a smallest element, which we denote to be n_{k+1}.
Observe that for every $\varepsilon > 0$ there is $k \in \mathbb{N}$ such that $|x_{n_k} - x_0| \leq \frac{4M}{2^k} < \varepsilon$. $\square$

Proposition 4.5. Every bounded, monotone sequence is convergent.

Proof. Assume WLOG that (x_n) is increasing, if not consider $(-x_n)$. By Bolzano-Weirstrass, (x_n) has a convergent subsequence (x_{n_k}) with limit x_0. Clearly, as x_{n_k} is increasing, $x_{n_k} \leq x_0$ for all k. Given $\varepsilon > 0$ we find $K \in \mathbb{N}$ such that $x_0 - x_{n_k} = |x_0 - x_{n_k}| < \varepsilon$ for all $k \geq K$.
By monotonicity $|x_0 - x_m| = |x_0 - x_m| \leq x_0 - x_{n_K} < \varepsilon$ for all $n \geq n_K$. $\square$

Definition 4.7 (Cauchy Sequence)**.** A sequence (x_n) is called a **cauchy sequence** if for each $\varepsilon > 0$ there is a $N \in \mathbb{N}$ such that $|x_n - x_m| < \varepsilon$ for $n, m \geq N$.

Theorem 4.4. A sequence is cauchy if and only if it is convergent.

Proof. If (x_n) is convergent with limit x_0, then for each $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that for $n, m \geq N$ $|x_n - x_m| \leq |x_n - x_0| + |x_m - x_0| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.
Now, recall that if a sequence is cauchy, then it is bounded. Then, by bolzano-weirstrass, there is a convergent subsequence (x_{n_k}) with limit x_0. Given $\varepsilon > 0$

we find $N \in \mathbb{N}$ such that $|x_n - x_m| < \frac{\varepsilon}{2}$ for $n, m \geq N$. Also, there is $K \in \mathbb{N}$ such that $K \geq N$ and $|x_{m_k} - x_0| < \frac{\varepsilon}{2}$ by convergence. Consequently, for $m \geq N$,

$$\begin{aligned} |x_m - x_0| &\leq |x_{m_k} - x_0| + |x_m - x_{m_K}| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Hence □

Lecture 7: Sequences (2), Limits, and Continuity

Tue 14 Sep 2021 13:01

Definition 4.8 (Limit Superior/Inferior)**.** A number L is called the **limit superior** of (x_n), denoted by $\limsup_{n\to\infty} x_n$, if for each $\varepsilon > 0$, the set $\{n \in \mathbb{N} : x_n > L + \varepsilon\}$ is finite and the set $\{m \in \mathbb{N} : x_m \geq L - \varepsilon\}$ is infinite. Similairly, the **limit inferior** of (x_n), denoted $\liminf_{n\to\infty} x_n$, is the number ℓ such that for each $\varepsilon > 0$ $\{n \in \mathbb{N} : x_n < \ell + \varepsilon\}$ is finite and $\{m \in \mathbb{N} : x_m \leq \ell + \varepsilon\}$

Remark. In order for the limit superior to exist, the sequence must be bounded from above. Furthermore, the limits superior and inferior are unique.

Proposition 4.6. Suppose (x_n) is bounded above, then $\limsup_{n\to\infty}(x_n)$ exists if and only if (x_n) has an accumulation point and if so, $\limsup_{n\to\infty}(x_n)$ is the largest accumulation point of (x_n).

Proof. Suppose $L = \limsup_{n\to\infty}(x_n)$ exists. Then for each $\varepsilon > 0$, the set $\{n \in \mathbb{N} : L - \varepsilon \leq x_m \leq L + \varepsilon\}$ is infinite hence L is an accumulation point.
Conversely, let A be the set of accumulation points of $\{x_n\}$. A is nonempty and has an upper bound. Hence, $L = \sup(A)$ exists. We wish to show $L = \limsup_{n\to\infty}(x_n)$. Since L is the least upper bound of A, for each $\varepsilon > 0$ there is a $K \in A$ such that $K > L - \frac{\varepsilon}{2}$. Since $K \in A$, we see for each $N \in \mathbb{N}$ there is a $n \geq N$ such that $|K - x_n| < \frac{\varepsilon}{2}$. Consequently, for such n, we have $x_n > K - \frac{\varepsilon}{2} > L - \varepsilon$. Hence,

$$|L - x_n| \leq |L - K| + |K - x_n| < \varepsilon.$$

Hence, $L \in A$ and the set $\{n \in \mathbb{N} : x_n > L - \varepsilon\}$ is infinite. Lastly, for every $\varepsilon > 0$, the set $\{n \in \mathbb{N} : x_n > L + \varepsilon\}$ is finite for otherwise by Bolzano-Weirstrass, there would be an accumulation point larger than L contradicting the fact that $L = \sup(A)$. □

Remark. (x_m) converges if and only if $\limsup_{m\to\infty} x_m$ and $\liminf_{m\to\infty x_m}$ both exist and are equal. This further guarantees $\lim\limits_{m\to\infty} x_m = \limsup_{m\to\infty} x_m = \liminf_{m\to\infty} x_m$.

Proposition 4.7. Suppose (x_n) has a limit superior. Then,

$$\limsup_{n\to\infty} x_n = \inf\{\sup\{x_k : k \in \mathbb{N} \text{ and } k \geq m\} : m \in \mathbb{N}\}.$$

Definition 4.9 (Series)**.** Let (x_n) to be a sequence, then the sequence $\left(\sum_{i=1}^{m} x_i\right)_m$ is called a **series**. We often abbreviate this $\sum_{i=1}^{\infty} x_k$ or $\sum_k x_k$. We also sometimes denote the limit of the series as $\sum_{i=1}^{\infty} x_i$.
The sum of a doubly infinite sequence $(x_k)_{k\in\mathbb{Z}}$ of nonnegative numbers is defined by

$$\sum_{k=-\infty}^{\infty} x_k = \sum_{k\in\mathbb{Z}} x_k = \sup\{\sum_{k=n}^{m} x_k : n, m \in \mathbb{Z}, m \le n\}..$$

Proposition 4.8.
- If $\sum_{i=1}^{\infty} x_i$ is convergent, then $\lim_{n\to\infty} x_n = 0$.
- If $\sum_{i=1}^{\infty} |x_i|$ is convergent, then $\sum_{i=1}^{\infty} x_i$ is convergent.

5 Continuity

Definition 5.1 (Continuity)**.** A function $f : S \to \mathbb{R}$ is **continuous at** x_0 if for each $\varepsilon > 0$ there is a $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ if $|x - x_0| < \delta$.
A function $f : S \to \mathbb{R}$ is **continuous on** X if for every $x_0 \in X$, f is continuous at x_0.
A function $f : S \to \mathbb{R}$ is **continuous** if every $x \in R$ has f being continous at x. A function $f : S \to \mathbb{R}$ is **uniformly continuous** on if for each $\varepsilon > 0$ there is a $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all $x, y \in S$ with $|x - y| < \delta$.
A function $f : S \to \mathbb{R}$ is **lipschitz continuous** if there exists $L \ge 0$ such taht $|f(x) - f(y)| \le L|x - y|$ for all $x, y \in S$.

Problem. Show lipschitz continuity implies uniform continuity implies continuity.

Definition 5.2. Let $S \subseteq \mathbb{R}$. A set $U \subseteq S$ is **relatively open**/**relatively closed** in S if there exists an open/closed set $V \subseteq \mathbb{R}$ such that $U = S \cap V$.

Proposition 5.1.
- Let $S \subseteq \mathbb{R}$. Then, $f : S \to \mathbb{R}$ is continuous at $x_0 \in S$ if and only if for every converget sequence (x_n) in S with $\lim_{n\to\infty} x_n = x_0$ we have $(f(x_n))$ is convergent with $\lim_{n\to\infty} f(x_n) = f(x_0)$.
- A function $f : S \to \mathbb{R}$ is continuous if and only if $f^{-1}(U)$ is relatively open in S for every open set $U \subseteq \mathbb{R}$.

Theorem 5.1 (Max-Min Value Theorem)**.** Suppose $f : S \to \mathbb{R}$ is continuous, with S being nonempty and compact. Then, f takes a maximum and a minimum value.

The proof of this is pretty straghtforward so we will not write it in detail here. Essentially one takes an open covering of $f(S)$ and notes that the preimage of this cover is also an open cover of S, hence there is a finite subcover which can be pushed back to a finite subcover of $f(S)$, As $f(S)$ is compact we have that it is closed and bounded, hence maximum and minimum values exist.

Theorem 5.2 (Intermediate Value Theorem)**.** Suppose $f : I \to \mathbb{R}$ is continuous on the interval I. Then, $f(I)$ is an interval.

Again, this theorem makes use of the fact that f is a continuous map and continuity preserves connectedness, hence I being an interval (connected) implies $f(I)$ is an interval (connected).

Definition 5.3 (Monotonicity)**.** A function $f : S \to \mathbb{R}$ is **increasing** if $x, y \in S$ with $x < y$ implies $f(x) \leq f(y)$. It is **strictly increasing** if the inequality is strict.
A function $f : S \to \mathbb{R}$ is **decreasing** if $x, y \in S$ with $x < y$ implies $f(x) \geq f(y)$. Again, it is **strictly decreasing** if the inequality is strict.
A function which is either increasing or decreasing is **monotone**.

Proposition 5.2. A monotone function $f : I \to \mathbb{R}$ on an interval I is continuous if and only if $f(I)$ is also an interval.

Lecture 8: Continuity (2) and Extended $\mathbb{R}$

Thu 16 Sep 2021 13:02

We begin with some results on continuity over intervals and inverses.

Recall. If $f : I \to \mathbb{R}$ is monotone with I being an interval. Then, f is continuous if and only if $f(I)$ is an interval.

Corollary 5. A continuous strictly monotone function $f : I \to \mathbb{R}$, with I being an interval, has a continuous inverse $f^{-1} : f(I) \to \mathbb{R}$.

Proposition 5.3. A strictly monotone function $f : I \to \mathbb{R}$, with I being an interval, has a continuous inverse.

Theorem 5.3 (Heine's Theorem)**.** A continuous function $F : S \to \mathbb{R}$ with S being compact is uniformly continuous.

Proof. For $\varepsilon > 0$ and $x \in S$, there is a $\delta_x > 0$ such that $|f(x) - f(y)| < \frac{\varepsilon}{2}$ if $|x - y| < \delta_x$. Let $U_x = \left(x - \frac{\delta_x}{2}, x + \frac{\delta_x}{2}\right)$. Since $\{U_x : x \in S\}$ is an open cover of S, there are $x_1, x_2, \ldots, x_n$ such that $\{U_{x_k} : 1 \leq k \leq n\}$ is a finite open subcover of S. Let $\delta = \min\{\frac{1}{2}\delta_{x_1}, \frac{1}{2}\delta_{x_2}, \ldots, \frac{1}{2}\delta_{x_n}\}$ and suppose $x, y \in S$ such that $|x - y| < \delta$. Then there is x_k for some $1 \leq k \leq n$ such that $x \in U_{x_k}$ and

$|x_k - y| \le |x - y| + |x_k - x| < \delta + \frac{\delta_{x_k}}{2} \le \delta_{x_k}$.
Consequently

$$|f(x) - f(y)| \le |f(x) - f(x_k)| + |f(x_k) - f(y)| < \varepsilon.$$

□

Justification for continuity

There is an equivalence between open sets/continuity and measurable sets/measurableness.

Definition 5.4 (Convergence of functions)**.** Let (f_n) be a sequence of functions $f_n : S \to \mathbb{R}$. Then

1. (f_n) **converges pointwise** if $(f_n(x))$ is convergent for every $x \in S$. The limit is defined pointwise for every $x \in S$ with $f(x) = \lim_{n\to\infty} f_n(x)$ with $f : S \to \mathbb{R}$ being a function.
2. (f_n) **converges uniformly** to the function $f : S \to \mathbb{R}$ if for each $\varepsilon > 0$, there $N \in \mathbb{N}$ such that for all $x \in S$, $|f(x) - f_n(s)| < \varepsilon$ if $n \ge N$.

Theorem 5.4. Suppose (f_n) is a sequence of continuous functions $f_n : S \to \mathbb{R}$ which converges uniformly to $f : S \to \mathbb{R}$. Then, f is continuous.

Proof. Let $x \in S$ and $\varepsilon > 0$. Then, there is $k \in \mathbb{N}$ such that $|f(y) - f_k(y)| < \frac{\varepsilon}{3}$ for all $y \in S$. Consequently, for any $y \in S$

$$\begin{aligned} |f(x) - f(y)| &\le |f(x) - f_k(x)| + |f_k(x) - f_k(y)| + |f(y) - f_k(y)| \\ &< \frac{2\varepsilon}{3} + |f_k(x) - f_k(y)|. \end{aligned}$$

Since f_k is continuous, we can pick a sufficient $\delta > 0$ such that this completes the proof. □

6 Extended ℝ

Recall that many objects such as the lim sup and lim inf required a boundedness assumption. We wish to discard this assumption when possible. Hence we introduce the following system.

Definition 6.1 (Extending Functions)**.** A function $h : \mathbb{R} \to \mathbb{R}$ is **extending** if h is strictly inreasing and $h(\mathbb{R}) = (-1, 1)$. Note that every extending function is continuous by these assumptions and has a continuous inverse.

Now, we introduce two external elements $-\infty, +\infty$ and we define $\infty = H^{-1}(1)$ and $-\infty = H^{-1}(-1)$ and we extend the ordering $\le$ such that $-\infty < \infty$ and $-\infty < x < \infty$ for every $x \in \mathbb{R}$.

Definition 6.2 (Extended Real Numbers)**.** We denote $R \cup \{-\infty, \infty\} = \overline{\mathbb{R}} = [-\infty, \infty]$ to be the **extended real numbers** for use with extending functions.

In this way, the extending function h extends from R to $\overline{R}$ and it retains its strictly increasing and the image requirement $h\left(\overline{\mathbb{R}}\right) = [-1, 1]$.

Notation.
- $(a, \infty] = \{x \in \overline{\mathbb{R}} : x > a\}$
- $[-\infty, a] = \{x \in \overline{\mathbb{R}} : x \leq a\}$
- $(a, \infty)_{\overline{\mathbb{R}}} = \{x \in \overline{\mathbb{R}} : a < x < \infty\}$
- and so on.

It is of note that the interval (a, ∞) is $\mathbb{R}$ is still defined as normal, it is only when it is chosen as part of $\overline{\mathbb{R}}$.
Now we examine the topology on $\overline{\mathbb{R}}$.

Definition 6.3 (Topology on $\overline{\mathbb{R}}$)**.**
1. $S \subseteq \overline{\mathbb{R}}$ is open/closed if $H(S)$ is relatively open/closed in $[-1, 1]$ for any extending function H.
2. $S \subseteq \overline{\mathbb{R}}$ has $\sup(S) = H^{-1}(\sup(H(S)))$.
3. A sequence (x_n) in $\overline{\mathbb{R}}$ is convergent (in $\overline{\mathbb{R}}$) if for any extending function H, $(H(x_n))$ isconvergent. In this case we define
$$\lim_{n\to\infty} x_n = H^{-1}\left(\lim_{n\to\infty} H(x_n)\right)$$
.
4. A point $x_0 \in \overline{\mathbb{R}}$ is an accumulation or cluster point of the sequence (x_n) in $\overline{\mathbb{R}}$ if for any extending function H we have $H(x_0)$ is an accumulation point of $(H(x_n))$.
5. Let (x_n) in $\overline{\mathbb{R}}$. Then,
$$\limsup_{n\to\infty} x_n = H^{-1}\left(\limsup_{n\to\infty} H(x_n)\right)$$
$$\liminf_{n\to\infty} x_n = H^{-1}\left(\liminf_{n\to\infty} H(x_n)\right)$$

Example.
- $\overline{\mathbb{R}}$ is open and closed in $\overline{\mathbb{R}}$.
- $\mathbb{R}$ is open but not closed in $\overline{\mathbb{R}}$.
- $(7, \infty] \mapsto (H(7), 1]$, hence it is open.

◇

Proposition 6.1. If (x_n) is a sequence with $x_n \in \overline{\mathbb{R}}$. Then

$$\limsup_{n\to\infty} x_n, \liminf_{x\to\infty} x_n \in \overline{\mathbb{R}}$$

with

$$\begin{aligned}\limsup_{n\to\infty} x_n &= \inf\left(\sup\{x_k : k \in \mathbb{N}, k \geq n\} : n \in \mathbb{N}\right)\\ &= \lim_{n\to\infty} \sup\{x_k : k \in \mathbb{N}, k \geq n\} \text{ and}\\ \liminf_{n\to\infty} x_n &= \sup\left(\inf\{x_k : k \in \mathbb{N}, k \leq n\} : n \in \mathbb{N}\right)\\ &= \lim_{n\to\infty} \inf\{x_k k \in \mathbb{N}, k \leq n\}\end{aligned}$$

Remark. A sequence (x_n) in $\mathbb{R}$ is said to converge to ∞ if it is convergent in $\overline{R}$ with $\lim\limits_{n\to\infty} x_n = \infty$.

Definition 6.4.

1. If $a \in (-\infty, \infty]$, then $a + \infty = \infty + a = \infty$.
2. If $a \in [-\infty, \infty)$ then $a + (-\infty) = (-\infty) + a = -\infty$.
3. If $a \in (0, \infty]$ then $a \cdot \infty = \infty \cdot a = \infty$.
4. If $a \in [-\infty, a)$ then $a \cdot \infty = \infty \cdot a = -\infty$.
5. If $a \in (-\infty, \infty) \setminus \{0\}$ then $\frac{\infty}{a} = \frac{1}{a} \cdot \infty$.
6. If $a \in (-\infty, \infty)$ then $\frac{a}{\infty} = \frac{a}{-\infty} = 0$.
7. If $a \in [-\infty, \infty] \setminus \{0\}$ then $\left|\frac{a}{0}\right| = \infty$ (though $\frac{a}{0}$ is left undefined).
8. $|\infty| = |-\infty| = \infty$ and $\infty^p = \infty$, $\infty^{-p} = 0$ for $p > 0$.
9. $0 \cdot \infty = \infty \cdot 0 = 0 \cdot (-\infty) = (-\infty) \cdot 0 := 0$.
10. $\frac{\infty}{\infty} = \frac{-\infty}{\infty} = \frac{\infty}{-\infty} = \frac{-\infty}{-\infty} := 0$

These last definitions go against our conventional logic involving ∞, but they are simply definitions which will be useful for measure theoretic results later on.

These conventions do have the unfortunate consequence that $\lim\limits_{n\to\infty} \frac{x_n}{y_n} \neq \frac{\lim_{n\to\infty} x_n}{\lim_{n\to\infty} y_n}$ in general for sequences $(x_n), (y_n), \left(\frac{x_n}{y_n}\right)$ in $\overline{\mathbb{R}}$. These facts still hold in sequences which converge in $\mathbb{R}$ (in $\overline{\mathbb{R}}$), it is simply when a sequence converges only in $\overline{\mathbb{R}}$ for which we have issues.

Remark. We left undefined $\infty - \infty$, $-\infty + \infty$, and $\frac{x}{0}$ for $x \in \overline{\mathbb{R}}$. Furthermore, we have $\frac{x}{y} = x \cdot \frac{1}{y}$ only if $x \in \overline{R}$, $y \in \overline{\mathbb{R}} \setminus \{0\}$.

Lecture 9: Extended ℝ (2) and Intro to Measure Theory

Tue 21 Sep 2021 12:20

Definition 6.5. Let $S \subseteq \mathbb{R}$ and $f : S \to \overline{\mathbb{R}}$. Then, we say f is continuous at $x_0 \in S$ if $H \circ f$ is continuous at x_0 on S for any extending function H. Similarity, we say f is continuous on S if $H \circ f$ is continuous on S for any extending function H.
Furthermore, we say f is (strictly) increasing/decreasing/monotone if $H \circ f$ is (strictly) increasing/decreasing/monotone.
Again, if (f_n) is a series of functions $f_n : S \to \overline{\mathbb{R}}$, we say (f_n) converges pointwise/uniformly to $f : S \to \overline{\mathbb{R}}$ if $(H \circ f_n)$ converges pointwise/uniformly to $H \circ f$ for any extending function H.

Definition 6.6. Let $S \subseteq \overline{\mathbb{R}}$ and suppose $a \in \overline{R}$ is an accumulation point of a sequence taking values in $S \setminus \{a\}$.
Then, a function $f : S \setminus \{a\} \to \overline{R}$ is said to have the limit $L \in \overline{\mathbb{R}}$ (relative to S) if for any extending function H and for each $\varepsilon > 0$ we have an $\delta > 0$ such that

$$|H(f(x)) - H(L)| < \varepsilon \text{ for all } x \in S \setminus \{a\} \text{ with } |H(x) - H(a)| < \delta.$$

We denote this by $\lim\limits_{x \to a} f(x) = L$ or $\lim\limits_{\substack{x \to a \\ S}} f(x) = L$

7 Measure Theory

Definition 7.1 (Length)**.** Let $I = (a, b)$ be an interval, then we define the measure function $\ell : \mathscr{P}(\mathbb{R}) \to \mathbb{R}_0^+$ with the following properties:

$$\begin{aligned} \ell(\varnothing) &= 0 \\ \ell(I) &= b - a, a, b \in \mathbb{R}. \end{aligned}$$

In all other cases $\ell(I = \infty)$.

We would like to generalize this notion by constructing a set function λ such that

$$\begin{aligned} \lambda : \mathscr{P}(\mathbb{R}) &\to [0, \infty] \\ \lambda(I) &= \ell(I) \text{ for intervals } I \subseteq \mathbb{R} \\ \lambda(x + S) &= \lambda(S) \text{ for } x \in \mathbb{R}, S \subseteq \mathbb{R}, x + S = \{x + s : s \in S\} \\ \text{if } \{S_m : m \in \mathbb{N}\} &\text{ is a countable disjoint collection of sets in } \mathbb{R}, \text{ then} \\ \lambda\left(\bigcup_{n=1}^{\infty} S_m\right) &= \sum_{n=1}^{\infty} \lambda(S_n) \\ &. \end{aligned}$$

It turns out such a function produces contradictions, hence it is poorly posed. Hence, we must alter or remove one of these constraints and as all of the properties are very straight forward it is best to alter the domain of λ itself.

Definition 7.2 (Measure)**.** Let $\mathscr{A}$ be a σ-algebra.

1. A set function $\mu : \mathscr{A} \to [0, \infty]$ is called **countably additive** if for every countable disjoint collection $\{S_n \in \mathscr{A} : n \in \mathbb{N}\}$ we have
$$\mu\left(\bigcup_{n\in\mathbb{N}} S_n\right) = \sum_{i=1}^{\infty} \mu(S_i).$$
2. A countable additive set function $\mu : \mathscr{A} \to [0, \infty]$ such that $\mu(\varnothing) = 0$ is called a **measure**.

Proposition 7.1. Let $\mu : \mathscr{A} \to [0, \infty]$. Then, μ is monotone in the sense that if $A, B \in \mathscr{A}$ with $A \subseteq B$, then we have $\mu(A) \leq \mu(B)$.

Proof. Since $B = A \cup (B \setminus A)$ and since μ is countably additive, then
$$\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A).$$
□

Now, we wish to extend our notion to arbitrary subsets of $\mathbb{R}$.

Notation. For $A \in \mathscr{P}(\mathbb{R})$, then $J(A)$ is defined to be the collection of all countable covers $\{I_n : n \in \mathbb{N}\}$ of A consisting of open, bounded intervals I_n.

Definition 7.3 (Lebesque Outer Measure)**.** Let $A \in \mathscr{P}(\mathbb{R})$, then the quantity $\mu^*(A) \in [0, \infty]$ is defined by
$$\mu^*(A) = \inf\{\sum_{i=1}^{\infty} \ell(J_i) : \{J_i : i \in \mathbb{N}\} \in J(A)\}.$$
This function $\mu^* : \mathscr{P}(\mathbb{R}) \to [0, \infty]$ is called the **Lebesque outer measure**.

Lemma 7.1.
1. The outer measure is monotone
2. The outer measure is translation invariant.
3. The outer measure is countable subadditive, that being for $\{S_n : n \in \mathbb{N}\}$ is a countable collection of sets, then $\mu^*\left(\bigcup_{n\in\mathbb{N}} S_n\right) \leq \sum_{n=1}^{\infty} \mu^*(S_n)$.

Proof.
1. Note that $J(A) \subseteq J(B)$, hence $\mu^*(A) \leq \mu^*(B)$.
2. Similarly, as each $\ell(J_i)$ is translationally invariant, we see μ^* is translationally invariant.

3. Let $\varepsilon > 0$. Then for each $n \in \mathbb{N}$, let $\{I_{n,k} : k \in \mathbb{N}\} \in J(S_n)$ be a collection of intervals such that $\sum_{k=1}^{\infty} \ell(J_{n,k}) \leq \mu^*(S_n) + \frac{\varepsilon}{2^n}$.
   Since, $\{I_{n,k} : n, k \in \mathbb{N}\} \in J\left(\bigcup_{n\in\mathbb{N}} S_n\right)$, we must have that

$$\begin{aligned}\mu^*\left(\bigcup_{n\in\mathbb{N}} S_n\right) &\leq \sum_{n\in\mathbb{N}}\sum_{k\in\mathbb{N}} \ell(I_{n,k}) \\ &\leq \sum_{n=1}^{\infty} \mu^*(S_n) + \varepsilon.\end{aligned}$$

   Since this holds for all $\varepsilon > 0$, this completes the proof.

□

Lemma 7.2. For every interval $I \subseteq \mathbb{R}$, the outer measure is $\mu^*(I) = \ell(I)$.

Proof. Let $I \subseteq \mathbb{R}$ be nonempty (if $I = \varnothing$ it is trivial that $\mu^*(\varnothing) = 0$). First, assume $I = [a, b]$ with $a \leq b \in \mathbb{R}$. Let $\{J_n : n \in \mathbb{N}\} \in J(I)$, then by Heine-Borel there is a finite subcovering $\{I_n : 1 \leq n \leq N\}$ such that no $I_n = \varnothing$. Note that as we will be taking the infimum, then any infinite collection containing this finite collection will be larger (or equal) hence will not matter in the infimum. Furthermore, we can assume that no interval J_n has $J_n \subseteq J_m$ for some $n \neq m$, and we can assume $I_n = (a_n, b_n)$ to be ordered such that $a_n < a_{n+1}$ for $1 \leq n \leq N-1$. Consequently, $b_n > a_{n+1}$ for $1 \leq n \leq N-1$ as otherwise their would be a gap in the covering, and $b_n > b$, as $b \in I$, and $a_1 < a$ by the same reasoning. Hence, we have an overlapping covering of $[a, b]$ by open bounded intervals (a_n, b_n). Hence,

$$\begin{aligned}\ell(J) &= b - a \\ &\leq b_N - a_1 \\ &\leq \sum_{i=1}^{N} (b_i - a_i) \\ &= \sum_{i=1}^{N} \ell(I_i) \\ &\leq \sum_{i=1}^{\infty} \ell(I_i) \\ &\Rightarrow \ell(I) \leq \mu^*(I).\end{aligned}$$

Now, we look to obtain the opposite inequality. Let $\varepsilon > 0$, then $\{(a - \varepsilon, b + \varepsilon)\} \in J(I)$, hence

$$\begin{aligned}\mu^*(I) &\leq b - a + 2\varepsilon \\ &= \ell(I) + 2\varepsilon\end{aligned}$$

as ε is arbitrary, we then have

$$\mu^*(I) \leq \ell(I).$$

Hence $\mu^*(I) = \ell(I)$ for this case.
Now, assume $I \in \{(a,b) : [a,b), (a,b]\}$ is any bounded interval with $a < b$. By monotonicity, for every $\varepsilon > 0$, we have

$$\begin{aligned}
\ell(I) - 2\varepsilon &= b - a - 2\varepsilon \\
&\leq \ell([a+\varepsilon, b-\varepsilon]) \\
&= \mu^*([a+\varepsilon, b-\varepsilon]) \\
&\leq \mu^*(I) \\
&\leq \mu^*(I) \\
&\leq \mu^*([a,b]) \\
&= b - a \\
&= \ell(I).
\end{aligned}$$

Hence, for every $\varepsilon > 0$, $\ell(I) - 2\varepsilon \leq \mu^*(I) \leq \ell(I)$, hence as ε is arbitrary $\mu^*(I) = \ell(I) = b - a$. This covers all the bounded cases, hence only the unbounded case remains.
If I is unbounded and $a \in I$, then $[a, a+n] \subseteq J$ for all $n \in \mathbb{N}$ or $[a-n, a] \subseteq J$ for all $n \in \mathbb{N}$. In either case, by the monotonicity of the outer measure, $\mu^*(I) \geq n$ for all $n \in \mathbb{N}$, hence $\mu^*(I) = \infty = \ell(I)$. This completes the proof. □

Hence, we have that μ^* conforms to all of our desired properties with the notable exception of countable additivity. This, of course, means μ^* is not in fact a measure, so we will again modify our measure function in order to induce a countably additive measure. This construction will come next lecture and will consist of again restricting the domain to a subset of $\mathscr{P}(\mathbb{R})$, the Lebesque measurable sets, a collection which will be introduced and formalized next lecture.

Lecture 10: Measure Theory (2) and Lebesque Measure

Thu 23 Sep 2021 12:58

Definition 7.4. A set $S \subseteq R$ is **measureable/Lebesque mesaurable** if for every $A \subseteq \mathbb{R}$,

$$\mu^*(A) = \mu^*(A \cap S) + \mu^*(A \cap S^c).$$

It actually suffices to show only

$$\mu^*(A) \geq \mu^*(A \cap S) + \mu^*(A \cap S^c).$$

Proposition 7.2. Every set $S \subseteq \mathbb{R}$ with $\mu^*(S) = 0$ is measurable.

Proof. For every $A \subseteq \mathbb{R}$, $\mu^*(A \cap S) \leq \mu^*(S) = 0$. Similairly, $\mu^*(A \cap S^c) = 0$. □

Definition 7.5. A set $S \subseteq \mathbb{R}$ with $\mu^*(S) = 0$ is said to have measure 0.

Lemma 7.3. For each $a \in \mathbb{R}$, (a, ∞) is measurable.

. Given $A \subseteq \mathbb{R}$ and $\varepsilon > 0$, we fine $\{I_n : n \in N\} \in J(A)$ such that

$$\mu^*(A) \geq \sum_{n=1}^{\infty} \ell(I_n) - \varepsilon.$$

Since $A \cap (a, \infty) \subseteq \bigcup_{n \in \mathbb{N}} (I_n \cap (a, \infty))$ and

$$A \cap (a, \infty)^c \subseteq \left(\bigcup_{n \in \mathbb{N}} (I_n \cap (-\infty, a)) \right) \cup (a - \varepsilon, a + \varepsilon).$$

It follows that $\mu^*(A \cap (a, \infty)) \leq \sum_{n=1}^{\infty} \ell(I_n \cap (a, \infty))$ and $\mu^*(A \cap (a, \infty)^c) \leq \sum_{n=1}^{\infty} \ell(I_n \cap (-\infty, a)) + 2\varepsilon$. As $\ell(I_n) = \ell(I_n \cap (a, \infty)) + \ell(I_n \cap (-\infty, a))$ as the singular point a will not change the length. Hence,

$$\begin{aligned}\mu^*(A) &\geq \sum_{n=1}^{\infty} \ell(I_n \cap (a, \infty)) + \sum_{n=1}^{\infty} \ell(I_n \cap (-\infty), a) - \varepsilon \\ &\geq \mu^*(A \cap (a, \infty)) + \mu^*(A \cap (a, \infty)^c) - 3\varepsilon.\end{aligned}$$

□

Proposition 7.3. The collection of Lebesque measurable sets in $\mathbb{R}$ is a σ-algebra $\mathscr{L}$ containing all Borel sets.

Proof. If the measurable sets form of σ-algebra $\mathscr{L}$, then $\mathscr{L}$ must contain all open and closed subsets of $\mathbb{R}$, since it contains all intervals of the form (a, ∞). To show that the measurable sets form a σ-algebra $\mathscr{L}$ we first note that $(a, a) = \varnothing$ and the complement of each measurable set are both measurable sets. This is due to the symmetry in the definition of measurbility

$$\mu^*(A) \geq \mu^*(A \cap S) + \mu^*(A \cap S^c).$$

Now, suppose $\{S_n : n \in \cap\}$ is a countable collection of measurable sets. Let $S = \bigcup_{n \in \mathbb{N}} S_n$, then we need only show S is measurable.
Given $A \subseteq \mathbb{R}$, we define a sequence with $A_1 = A$, $A_{n+1} = A \cap (\bigcap_{k=1}^{n} S_k^c)$. Hence, $A_2 = A \cap S_1^c$, $A_3 = A \cap (S_1^c \cap S_2^c)$. Now, note that $A_{n+1} = A_n \cap S_n^c$, hence the sequence is decreasing in size. And $A \cap S = \bigcup_{k \in \mathbb{N}} (A_k \cap S_k)$. We present a short proof of this claim.
Note that for $x \in A \cap S$, there is a smallest positive integer k such that $x \in S_k$. If $k = 1$, then $x \in A_1 \cap S_1$, if $k > 1$, then $x \notin S_n$ for any $n < k$, consequently $x \in A_k$ by construction. Hence, $x \in A_k \cap S_k$, so $A \cap S \subseteq \bigcup_{k \in \mathbb{N}} (A_k \cap S_k)$.
Now, $\bigcup_{k \in \mathbb{N}} (A_k \cap S_k) \subseteq A \cap S$, as each $A_k \in A$ and $S_k \in S$, hence their intersection and subsequent union are also contained. Hence the equality is shown

$$A \cap S = \bigcup_{k \in \mathbb{N}} (A_k \cap S_k).$$

By measurability of S_n, we know any set A has

$$\mu^*(A_n) = \mu^*(A_n \cap S_n) + \underbrace{\mu^*(A_n \cap S_n^c)}_{A_{n+1}}$$

.
Hence, by induction, we have $\mu^*(A) = \mu^*(A_1) = \sum_{k=1}^n \mu^*(A_k \cap S_k) + \mu^*(A_{n+1})$.
Since $A \cap \left(\bigcap_{k\in\mathbb{N}} S_k^c\right) = A \cap S^c \subseteq A_{n+1}$ for any n.
Hence,

$$\mu^*(A) \geq \sum_{k=1}^n \mu^*(A_i \cap S_i) + \mu^*(A \cap S^c).$$

Finally, as $\bigcup_{k\in\mathbb{N}} (A_k \cap S_k) = A \cap S$ and since μ^* is contably subadditive, we obtain

$$\begin{aligned}
\mu^*(A) &\geq \sum_{k=1}^\infty \mu^*(A_k \cap S_k) + \mu^*(A \cap S^c) \\
&\geq \mu^*\left(\bigcup_{k\in\mathbb{N}} (A_k \cap S_k)\right) + \mu^*(A \cap S^c) \\
&= \mu^*(A \cap S) + \mu^*(A \cap S^c).
\end{aligned}$$

□

Definition 7.6 (Lebesque Measure)**.** The **Lebseque Measure** of a measurable set $S \subseteq \mathbb{R}$, denoted by $\mu^*(S)$ is defined by $\mu(S) = \mu^*(S)$. The set function $\mu : \mathscr{L} \to [0, \infty]$ is called the **Lebesque Measure**.

Theorem 7.1. The Lebesque measure μ is a measure on $\mathscr{L}$ such that

- $\mu(I) = \ell(I)$ for every interval $I \subseteq \mathbb{R}$.
- μ is translation invariant.
- μ is countablly additive.

Proof.

1. As μ^* has the interval property, μ trivially inherits this,
2. Similarly, as μ^* was translationaly invariant, we see μ inherits this.
3. Let $\{S_k : k \in \mathbb{N}\}$ be a countable, disjoint collection of measurable sets and define $T_n = \bigcup_{k=n}^\infty S_k$ for $n \in \mathbb{N}$.
   Since, $T_{n+1} = T_n \cap S_n^c$ we have

   $$\mu(T_n) = \mu(T_n \cap S_n) + \mu\left(\underbrace{T_n + S_n^c}_{=T_{n+1}}\right)$$

   by measurability of S_n.

   Consequently, $\mu(T_1) = \sum_{k=1}^n \mu\left(\underbrace{T_k \cap S_k}_{=S_n}\right) + \mu(T_{n+1}) \geq \sum_{k=1}^n \mu(S_k)$ for every $n \in \mathbb{N}$. Thus $T_1 = \bigcup_{k\in\mathbb{N}} S_k$ gives $\mu\left(\bigcup_{k\in\mathbb{N}} S_k\right) \geq \sum_{k=1}^{n} \mu\left(S_k\right)$. And as we already know the inequality goes in the other direction by subadditivity of μ^*, we see equality holds.

□

Corollary 6. Every countable set of real numbers is measurable with measure 0.

Proof. Let C be our countable sets and note that $C = \bigcup_{k\in\mathbb{N}}\{x_k\}$ with $x_k \neq x_m$ for $k \neq m$. Then, we see that

$$\mu\left(\bigcup_{k\in\mathbb{N}}\{x_k\}\right) = \sum_{k=1}^{\infty} \mu\left(\{x_k\}\right) = 0.$$

□

Theorem 7.2 (Properties of Lebesque Measure)**.** Let $S \subseteq \mathbb{R}$, the following are equivalent

1. S is measurable.
2. For each $\varepsilon > 0$, there is an open set O and a closed set C such that $C \subseteq S \subseteq O$ and $\mu\left(O \setminus C\right) < \varepsilon$.
3. There is a G_δ set G and a F_σ set F such that $F \subseteq S \subseteq G$ and $\mu\left(G \setminus F\right) = 0$.
4. For each $\varepsilon > 0$, there are measurable sets G and F such that $F \subseteq S \subseteq G$ and $\mu\left(G \setminus F\right) < \varepsilon$.

We will prove this result next time, though it is completely trivial that $3 \Rightarrow 4$, so we will primarily focus on proving $1 \Rightarrow 2$ and $4 \Rightarrow 1$.

Lecture 11: Measure Theory (3)

Tue 28 Sep 2021 13:00

We prove the final theorem from last lecture.

Proof.
- $(1 \Rightarrow 2)$. There are 2 cases, S being bounded and S being unbounded.
  If S is bounded, there is an interval $(a, b) \supseteq S$, $a, b \in \mathbb{R}$. Then for any given $\varepsilon > 0$, we find $\{I_k : k \in \mathbb{N}\} \in J\left(S\right)$ and $\{J_k : k \in \mathbb{N}\} \in J\left([a, b] \setminus S\right)$ such that $\mu\left(s\right) \geq \sum_{k=1}^{\infty} \ell\left(I_k\right) - \frac{\varepsilon}{3}$ and $\mu\left([a, b] \setminus S\right) \geq \sum_{k=1}^{\infty} \ell\left(J_k\right) - \frac{\varepsilon}{3}$.
  Let $O = \bigcup_{k\in\mathbb{N}} I_k$, $U = \bigcup_{k\in\mathbb{N}} J_k$ and $C = [a, b] \setminus U$. Then, $C \subseteq S \subseteq O$.
  Note that O, U are open and C is closed. Then,

$$\begin{aligned} \mu\left(S\right) &\geq \mu\left(O\right) - \frac{\varepsilon}{3} \\ \mu\left([a, b] \setminus S\right) &\geq \mu\left(U\right) - \frac{\varepsilon}{3}.. \end{aligned}$$

Furthermore, U, C are disjoint and $\mu(U) < \infty$ (as it is an interval minus a measurable set) and $[a,b] \subseteq U \cup C$. Hence,

$$\begin{aligned} \mu(C) &\geq \mu([a,b]) \setminus \mu(U) \\ &= b - a - \mu(U). \end{aligned}$$

Then, since $\mu(C) \leq \mu(S) < \infty$, we have

$$\begin{aligned} \mu(O \setminus C) &= \mu(O) - \mu(C) \\ &\leq \frac{\varepsilon}{3} + \underbrace{\mu(S) - (b-a)}_{=-\mu([a,b\setminus S])} + \mu(U) \\ &= \frac{\varepsilon}{3} - \mu([a,b] \setminus S) + \mu(U) \\ &\leq \frac{2\varepsilon}{3} \\ &< \varepsilon. \end{aligned}$$

For a general S, let $S_n = S \cap [n, n+1]$, $n \in \mathbb{Z}$. Then, there are open O_n and closed C_n such that $C_n \subseteq S_n \subseteq O_n$ and $\mu(O_n \setminus C_n) < \frac{\varepsilon}{3 \cdot 2^{|n|}}$.
Let $O = \bigcup_{n\in\mathbb{Z}} O_n$ and $C = \bigcap_{n\in\mathbb{Z}} C_n$. Then, O is open and C is closed by definition and we see $O \setminus C = \bigcup_{n\in\mathbb{Z}} (O_n \setminus C_n)$ by demorgen and we have $C \subseteq S \subseteq O$. Then,

$$\begin{aligned} \mu(O \setminus C) &\leq \sum_{n\in\mathbb{Z}} \mu(O_n \setminus C_n) \\ &< \sum_{n\in\mathbb{Z}} \frac{\varepsilon}{3 \cdot 2^{|n|}} \\ &= \varepsilon \text{ by geometric summation.} \end{aligned}$$

- $(2 \Rightarrow 3)$. For each $n \in \mathbb{N}$, there are closed C_n and open O_n such that $C_n \subseteq S \subseteq O_n$ and $\mu(O_n \setminus C_n) < \frac{1}{n}$. Let $F = \bigcup_{n\in\mathbb{N}} C_n$ and $G = \bigcap_{O_n}$. Then, F is a F_σ set and G is a G_δ set. Then, we have $F \subseteq S \subseteq G$ and $\mu(G \setminus F) \leq \mu(O_n\ C_n) < \frac{1}{n}$ for all $n \in \mathbb{N}$. Hence, $\mu(G \setminus F) = 0$.
- $(3 \Rightarrow 4)$. This is immediately obvious as F_σ and G_δ sets are measurable.
- $(4 \Rightarrow 1)$. Let $A \subseteq \mathbb{R}$ and $\varepsilon > 0$. Then $S^c \subseteq G \cup (G \cap F^c)$. Then, $A \cap S^c \subseteq (A \cap G^c) \cup (G \cap F^c)$.
  Hence,

$$\begin{aligned} \mu^*(A \cap S^c) &\leq \mu^*(A \cap G^c) + \underbrace{\mu^*(G \cap F^c)}_{<\varepsilon} \\ &\leq \mu^*(A \cap G^c) + \varepsilon. \end{aligned}$$

  And, as G is measurable, we have

$$\mu^*(A) = \mu^*(A \cap G) + \mu^*(A \cap G^c) \geq \mu^*(A \cap S) + \mu^*(A \cap S^c) - \varepsilon$$

  . Hence, in the infimum we have

$$\mu^*(A) \geq \mu^*(A \cap S) + \mu^*(A \cap S^c).$$

So, S is measurable.

$\square$

Definition 7.7 (Nested Sets)**.** A countable collection of sets $\{S_k : k \in \mathbb{N}\}$ is called

1. **ascending** if $S_k \subseteq S_{k+1}$ for all k.
2. **descending** if $S_{k+1} \subseteq S_k$ for all k.

Lemma 7.4.

1. If $\{S_k : k \in \mathbb{N}\}$ is an ascending collection of measurable sets, then $\mu\left(\bigcup_{k\in\mathbb{N}} S_k\right) = \lim_{k\to\infty} \mu(S_k)$.
2. If $\{S_k : k \in \mathbb{N}\}$ is a descending collection of measurable sets and $\mu(S_1) < \infty$. Then, $\mu\left(\bigcap_{k\in\mathbb{N}} S_k\right) = \lim_{k\to\infty} \mu(S_k)$.

Proof.

1. It suffices to consider the case $\mu(S_k) < \infty$ for all k, else the union and limit both trivially have measure ∞. Define $S_0 = \varnothing$, $X_n = S_n \setminus S_{n-1}$. Then, $\{X_k : k \in \mathbb{N}\}$ is a disjoint collection of measurable sets such that $\bigcup_{k\in\mathbb{N}} X_k = \bigcup_{k\in\mathbb{N}} S_k$. Hence, as we know the lebesque measure to be countably additive, we have

$$\begin{aligned}
\mu\left(\bigcup_{k\in\mathbb{N}} X_k\right) &= \sum_{k=1}^{\infty} \mu(X_k) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} \mu(X_k) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} \left(\mu(S_k) - \mu(S_{k-1})\right) \\
&= \lim_{k\to\infty} \mu(S_k).
\end{aligned}$$

2. Let $X_n = S_1\ S_n$. Then, $\{X_k : k \in \mathbb{N}\}$ is an ascending collection of measurable sets such that $\bigcup_{k\in\mathbb{N}} X_k = S_1 \setminus \left(\bigcap_{k\in\mathbb{N}} S_k\right)$. Since $S_k \subseteq S_1$ and $\mu(S_1) < \infty$ we have by the first lemma that

$$\begin{aligned}
\mu(S_1) - \mu\left(\bigcap_{k\in\mathbb{N}} S_k\right) &= \mu\left(\bigcup_{k\in\mathbb{N}} X_k\right) \\
&= \lim_{k\to\infty} \mu(X_k) \\
&= \mu(S_1) - \lim_{k\to\infty} \mu(S_k).
\end{aligned}$$

As $\mu(S_1)$ is finite we know this to be well defined, hence

$$\mu\left(\bigcap_{k\in\mathbb{N}} S_k\right) = \lim_{k\to\infty} \mu(S_k).$$

□

Theorem 7.3 (Borel-Cantelli Lemma)**.** Suppose $\{S_k : k \in \mathbb{N}\}$ is a countable collection of measurable sets such that $\sum_{k=1}^{\infty} \mu(S_k) < \infty$. Then, the set of all $x \in \mathbb{R}$ which belong to an infinite subcollection of $\{S_k : k \in \mathbb{N}\}$ has measure 0.

Proof. Note that x belongs to an infinite subcollection of $\{S_k : k \in \mathbb{N}\}$ if and only if $x \in \bigcap_{k\in\mathbb{N}} \bigcup_{n=k}^{\infty} S_n$.
Then, the collection $\{\bigcup_{n=k}^{\infty} S_n : k \in \mathbb{N}\}$ is descending and $\mu\left(\bigcup_{n\in\mathbb{N}} S_n\right) \leq \sum_{n=1}^{\infty} \mu(S_n) < \infty$. Hence, by the preceding lemma, we have

$$\begin{aligned}
\mu\left(\bigcap_{k\in\mathbb{N}} \bigcup_{n=k}^{\infty} S_n\right) &= \lim_{k\to\infty} \mu\left(\bigcup_{n=k}^{\infty} S_n\right) \\
&\leq \lim_{k\to\infty} \sum_{n=k}^{\infty} \mu(S_n) \\
&= 0
\end{aligned}$$

.

This final equality is because for all $\varepsilon > 0$ there is a $K \in \mathbb{N}$ such that for $k \geq K$ we have

$$\left| \sum_{i=1}^{\infty} \mu(S_i) - \sum_{i=1}^{k-1} \mu(S_i) \right| < \varepsilon.$$

□

Problem. 1. Is every set measurable?

2. Is every set of measure 0 countable?

3. Is every measurable set Borel?

Lecture 12: Negative Results of Measure Theory

Tue 05 Oct 2021 13:02

Lecture 13: Negative Results (2) and Measurable Functions

Tue 05 Oct 2021 13:02

We construct a cantor set.
First, suppose the interval $[0,1]$ and a series of sets $C_0, C_1, \ldots$ where $C_i = C_{i-1} \setminus D_i$ where D_i is just the set consisting of the middle thirds of each interval of C_{i-1}. Then, we let $C = \bigcap_{k\in\mathbb{N}} C_k$. We then define the nth partition of $[0,1] \setminus C_k$ to be $J_{k,n}$. We define $\mathscr{O} = \bigcup_{k,n\in\mathbb{N}} J_{k,n}$ and $\xi : \mathscr{O} \to \mathbb{R},\ x \in J_{k,n} \mapsto \frac{n}{2^k}$. We see this is well defined by an inductive argument.

Definition 7.8 (Cantor-Lebesque Function)**.** We define

$$\varphi : [0,1] \longrightarrow \mathbb{R}$$
$$x \longmapsto \varphi(x) = \begin{cases} 0, & x = 0 \\ \xi(x), & x \in \mathscr{O} \\ \sup\{\xi(y) : y \in \mathscr{O} \cap [0,x)\}, & x \in C \setminus \{0\} \end{cases}$$

to be the **Cantor-Lebesque Function**

Proposition 7.4. φ is a continuous increasing function such that $\varphi([0,1]) = [0,1]$.

Proof. It is clear ξ is and this guarantees φ to be increasing.
Next, note $\varphi(0) = 0$ and $\varphi(1) = 1$. Hence, we have the intermediate value theorem guaranteeing the image is $[0,1]$ if φ is continuous.
We see φ is continuous on $\mathscr{O}$ since it is constant on each interval $J_{k,n}$. Now, we consider $x \in C \setminus \{0,1\}$. For a given ε, let $k \in \mathbb{N}$ such that $\frac{1}{2^k} < \varepsilon$. Then, there is $n \in \mathbb{N}$ such that $1 \leq n \leq 2^k - 2$ such that for all $u \in J_{k,n}$, $v \in J_{k,n+1}$ such that for all u, v we find $u < x < v$. Let $a_k \in J_{k,n}$ $b_k \in J_{k,n+1}$ then by monotinicity of φ, for all $y \in [0,1]$ with $|x - y| < \delta = \min\{x - a_k, x + b_k\}$ we find

$$\begin{aligned} |\varphi(x) - \varphi(y)| &\leq \varphi(b_k) - \varphi(a_k) \\ &= \frac{n+1}{2^k} - \frac{n}{2^k} \\ &= \frac{1}{2^k} \\ &< \varepsilon. \end{aligned}$$

Finally, given $\varepsilon > 0$, we take $k \in \mathbb{N}$ such that $\frac{1}{2^k} < \varepsilon$ and let $c_k \in I_{k,1}$, $d_k \in I_{k,2^k-1}$. Then, for $o \leq y \leq c_k$, we find

$$\begin{aligned} |\varphi(0) - \varphi(y)| &= |\varphi(y)| \\ &\leq \varphi(c_k) \\ &= \frac{1}{2^k} \\ &< \varepsilon. \end{aligned}$$

Similairly, for $d_k < y \leq 1$, we find

$$\begin{aligned} |\varphi(1) - \varphi(y)| &\leq |1 - \varphi(d_k)| \\ &= 1 - \frac{2^k - 1}{2^k} \\ &= \frac{1}{2^k} \\ &< \varepsilon. \end{aligned}$$

$\square$

Definition 7.9 (Modified Cantor-Lebesque Function)**.** Let $\psi = x + \varphi(x)$ be the **modified Cantor-Lebesque Function**. It is clear ψ is continuous, strictly increasing and has , $\psi([0,2]) = [0,2]$.

Proposition 7.5. The function ψ has the following properties

1. $\psi(C)$ is measurable with $\mu(\psi(C)) = 1$.
2. There is a measurable set $S \subseteq C$ such that $\psi(S)$ is not measurable.

Proof.

- Note that $[0,1] = C \cup \mathscr{O}$ and ψ is injective and continuous. Hence, we have $[0,2] = \psi(C) \cup \psi(\mathscr{O})$ with $\psi(C) \cap \psi(\mathscr{O}) = \varnothing$. Since ψ is strictly increasing, we know ψ^{-1} is well-defined and continuous. Hence, ψ is an open map and we see $\psi(\mathscr{O})$ is open in $[0,2]$, hence $\psi(C)$ is closed. Hence, both sets are measurable. We see $\psi(\mathscr{O})$ is the union of a countable collection of open disjoint intervals, $\{I_i : i \in \mathbb{N}\}$ such that $\varphi \mid J_i$ is constant by construction. Hence, we hve for each $i \in \mathbb{N}$ we find $\psi(I_n) = x_i + I_i$ where $x_i \in [0,1]$ is a constant. Since ψ is injective, we find it preserves disjointness, hence the collection $\{\psi(I_i) : i \in \mathbb{N}\}$ is disjoint.
  Then, by countable additivity and translation invariance of μ we find

$$\begin{aligned}
\mu(\psi(\mathscr{O})) &= \mu\left(\bigcup_{i \in \mathbb{N}} I_i\right) \\
&= \bigcup_{i \in \mathbb{N}} \psi(I_i) \\
&= \sum_{i=1}^{\infty} \mu(\psi(I_i)) \\
&= \sum_{i=1}^{\infty} \ell(x_i + I_i) \\
&= \sum_{i=1}^{\infty} \ell(I_i) \\
&= \mu(\mathscr{O}).
\end{aligned}$$

Since, $\mu(C) = 0$, we find

$$\mu(\mathscr{O}) = \mu([0,1] \setminus C) = \mu([0,1]) = 1.$$

Consequently, $\mu(\psi(\mathscr{O})) = 1 = \mu(\mathscr{O})$. Hence, we find $\mu(\psi(C)) = 1$.

Since $\psi(C)$ has positive measure, it contains a nonmeasurable subset T, however, we see $S = \psi^{-1}(T)$ is measurable as $S \subseteq C$ and $\mu(C) = 0$. $\square$

Corollary 7. There is a measurable set $S \subseteq C$ such that S is not borel.

Proof. Since ψ has a continuous inverse, we see it maps borel sets to borel sets. Let S be a subset of C such that $\psi(S)$ is not measurable. Since $\psi(S)$ is not measurable, it is not a borel set. Hence S is not borel, but it was measurable with measure 0. $\square$

8 Measurable Functions

Definition 8.1 (Measurable Functions)**.** A function $f : S \to \overline{\mathbb{R}}$ is **Lebesque-measurable** on S if $S \subseteq \mathbb{R}$ is measurable and $f^{-1}((c, \infty])$ is a measurable set for every $c \in \mathbb{R}$. This is equivalent to the condition that $f^{-1}(B)$ is measurable for all $B \in \overline{\mathscr{B}}$, the extended borel σ-algebra.

Proposition 8.1. Let $S \subseteq \mathbb{R}$ be measurable, then a function $f : S \to \overline{\mathbb{R}}$ is measurable if and only if one of the following holds for all $c \in \mathbb{R}$:

- $f^{-1}([c, \infty])$ is measurable,
- $f^{-1}([-\infty, c])$ is measurable,
- $f^{-1}([-\infty, c))$ is measurable.

Definition 8.2. The extended Borel σ-algebra, $\overline{\mathscr{B}}$ consists of all subsets $B \subseteq \overline{\mathbb{R}}$ such that $B \setminus \{-\infty, \infty\} \in \mathscr{B}$.

Remark. It is clear $\overline{\mathscr{B}}$ is the smallest σ-algebra containing all open subsets of $\overline{\mathbb{R}}$.

Lecture 14: Measurable Functions (2)

Thu 07 Oct 2021 12:58

Recall. A function $f : S \to \mathbb{R}$ was measurable if S is measurable and $f^{-1}((c, \infty])$ is measurable for all $c \in \mathbb{R}$. There was an equivalent definition using the extended borel σ-algebra that we will use occasionally.

Proposition 8.2. Suppose $f : S \to \overline{\mathbb{R}}$ is continuous on the measurable set S, then f is measurable.

Proof. Let H be an extending function, then we must show $H \circ f$ is continuous. We see any subray , $f(X_0) = (c, \infty]$ will have $(H \circ f)(X_0) = (\hat{c}, 1]$. We know the preimage of this to be open in S, hence measurable. $\square$

Proposition 8.3. Let $S \subseteq \mathbb{R}$. Suppose $f : S \to \mathbb{R}$ is measurable. and let $g : B \to \mathbb{R}$ with $B \in \overline{\mathscr{B}}$ and $f(S) \subseteq B$. Then, $g \circ f : S \to \mathbb{R}$ is measurable.

Proof. For $c \in \mathbb{R}$, we note that $(g \circ f)^{-1}((c, \infty]) = f^{-1}\left(g^{-1}((c, \infty])\right)$. By continuity of g, we know $g^{-1}((c, \infty]) \in \overline{\mathscr{B}}$. And, since f is measurable, we find

$f^{-1}\left(g^{-1}\left((c,\infty]\right)\right)$. ☐

Corollary 8. Let $S \subseteq \mathbb{R}$ and $f : S \to \mathbb{R}$ to be a measurable function. Then, for every $\alpha \in \mathbb{R}$ and $0 < \rho < \infty$, we find αf and $|f|^{\rho}$ are measurable.

Proof. We see the functions $g(u) = \alpha u$ on $\overline{\mathbb{R}}$ and $h(u) = |u|^{\rho}$ on $\overline{\mathbb{R}}$ to be the corresponding functions. We see the case h is clearly continuous and well defined. On the other hand g may be poorly defined if $\alpha = 0$ and $f(x) = \infty$. Recall, however, we had $0 \cdot \pm\infty = 0$ so g is just the zero functions and we see continuity holds. ☐

Definition 8.3 (Almost-everywhere)**.** Let S be measurable, then a property is said to hold true **almost everywhere** on S or **for almost all** $x \in S$ if there is a set T with $\mu(T) = 0$ and the property holds on all of $S \setminus T$.

Proposition 8.4. Let $S \subseteq \mathbb{R}$ and suppose $f, g : S \to \overline{\mathbb{R}}$ such that f is measurable and $g = f$ almost everywhere on S, then g is measurable.

Proof. Let $T = \{x \in S : f(x) \neq g(x)\}$. Fix $c \in \mathbb{R}$ and let $F = f^{-1}((c,\infty]) \setminus T$ and $G = f^{-1}((c,\infty]) \cup T$. Clearly, both F and G are measurable. Furthermore, $F \subseteq G$ and $\mu(G \setminus F) = \mu(T) = 0$. Since, $F \subseteq g^{-1}((c,\infty]) \subseteq G$. And, by an earlier characterization we recall that a set X is measurable if and only if there were nested sets around it with a difference of measure 0 . Hence, g is measurable. ☐

Remark. Suppose $f : S \to \overline{\mathbb{R}}$ is a measurable set and $S \subseteq X \subseteq \mathbb{R}$. If $\mu(X \setminus S) = 0$ and $h : X \to \overline{\mathbb{R}}$ is any extension of f , then h is measurable since $h^{-1}((c,\infty]) = f^{-1}((c,\infty]) \cup \{x \in X \setminus S : h(x) \in (c,\infty]\}$. This is the union of a measurable set with a set of measure 0, so we see h is measurable.

Notation. Instead of saying that every extension of a measurable function $f : S \to \overline{\mathbb{R}}$ to a function $h : X \to \overline{\mathbb{R}}$, we often just say f is measurable on X as long as it is defined almost everywhere on X and is measurable on that set.

Proposition 8.5. Suppose $f : I \to \overline{\mathbb{R}}$ is monotone on $I \subseteq \mathbb{R}$. Then, the set of all points in I where f fails to be continuous is countable, hence measure 0. Another characterization is that f is continuous almost everywhere, hence f is measurable.

Proof. It suffices to consider the case f is increasing and I open. Let E be the set of all $x \in I$ where f fails to be continuous. For $x \in E$ let $\alpha_x = \sup(\{f(z) : z < x\} z \in J)$ and $\beta_x = \inf(\{f(z) : z > x\} z \in J)$. Since f is not continuous at x, we find the interval $(\alpha_x, \beta_x) = I_x$ to be nonempty. Also, if $x, y \in E$ are distinct with $x < y$ we find $\beta_x <= \alpha_y$. Hence, we find $I_x \cap I_y = \varnothing$. Since each interval I_x for $x \in E$ contains a rational number, we see E is countable. Hence, $\mu(E) = 0$ and we see $f|_{I \setminus E}$ is continuous on $I \setminus E$ which is measurable, hence the restriction is measurable and as f coincides with its restriction almost everywhere, we see f is measurable. $\square$

Definition 8.4 (Finite Functions).

- Let $S \subseteq \mathbb{R}$. A function $f : S \to \overline{\mathbb{R}}$ is called **finite on** S if $|f(x)| < \infty$ for all $x \in S$.
- Let $f, g : S \to \overline{\mathbb{R}}$ Then we say $f < g$ if $f(x) < g(x)$ for all $x \in S$. Similarly for all other inequalities.
- f is called **nonnegative** if $f \geq 0$ and **positive** if the inequality is strict.

Proposition 8.6. Let $f, g : S \to \overline{\mathbb{R}}$ be measurable and finite almost everywhere. Then, $f + g, f - g, f \cdot g$ are measurable.
If $g(x) \neq 0$ for almost every $x \in S$, then $\frac{f}{g}$ is measurable.

Proof.

1. First, we prove addition. We may assume f, g are finite on S . Then, $h = f + g$ is well defined. Since for $x \in S$, we have $h(x) > q$ for $c \in R$ if and only if there is a $q \in \mathbb{Q}$ such that $f(x) > q$ and $g(x) > c - q$, we have

$$\begin{aligned} h^{-1}((c, \infty]) &= h^{-1}((c, \infty)) \text{ by finiteness.} \\ &= \bigcup_{q \in \mathbb{Q}} f^{-1}((q, \infty)) \cup g^{-1}(c - q, \infty). \end{aligned}$$

   Hence, h as measurable as these are all measurable sets. If f, g are measurable, then so are $f, -g$, hence $f + (-g) = f - g$

2. With addition, subtraction is completely trivial,

3. Now multiplication, Let h be any measurable finite function on S. Consider $(h)^2$. If $c \geq 0$, we have

$$\left((h)^2\right)^{-1}((c, \infty)) = h^{-1}\left((-\infty, \sqrt{c})\right) \cup h^{-1}\left((\sqrt{c}, \infty)\right).$$

   If $c < o$, then

$$\left((h)^2\right)^{-1}((c, \infty)) = h^{-1}(\mathbb{R}) = S.$$

   As in either case we had the preimage being measurable, we see $(h)^2$ is measurable. Since $f \cdot g = \frac{1}{2}(f + g)^2 - \frac{1}{2}(f)^2 - \frac{1}{2}\left(g^2\right)$ being the sum, constant multiple and square of measurable functions yields $f \cdot g$ to be measurable.

4. Lastly, let $h = \frac{1}{g}$, and note we can assume g is nonzero for all S , hence h is well defined on S and h is finite. If $c > 0$ we see $h^{-1}((c, \infty)) = g^{-1}(\left(0, \frac{1}{c}\right))$. As this interval is open and borel, we see $g^{-1}\left(\left(0, \frac{1}{c}\right)\right)$ is borel, hence $h^{-1}((c, \infty))$ is measurable.

Similairly, if $c = 0$, we see $h^{-1}((0,\infty)) = g^{-1}((0,\infty))$.
Lastly, if $c < 0$ we have $h^{-1}(c,\infty) = g^{-1}((-\infty,\frac{1}{c})) \cup g^{-1}((0,\infty)) = g^{-1}\left(\left[\frac{1}{c},0\right)^c\right)$ hence measurable. This completes the proof.

□

Lecture 15: Measurable Functions (3) and Simple Functions

Thu 14 Oct 2021 13:01

Proposition 8.7. Let (f_n) be a sequence of measurable functions $f_n : S \to \overline{\mathbb{R}}$. Then, we define $f, g, F, G : S \to \overline{\mathbb{R}}$ with

- $f(x) = \sup\{f_n(x) : n \in \mathbb{N}\}$,
- $g(x) = \inf\{f_n(x) : n \in \mathbb{N}\}$,
- $F(x) = \limsup_{n\to\infty} f_n(x)$,
- $G(x) = \liminf_{n\to\infty} f_n(x)$

all being measurable.

Proof.
- Note that $f(x) > c$ if and only if there is an n such that $f_n(x) > c$. Hence, $f^{-1}((c,\infty]) = \bigcup_{n\in\mathbb{N}} f_n^{-1}((c,\infty)$ is measurable.
- It it clear $g(x) = -\sup\{-f_n(x) : n \in \mathbb{N}\}$.
- Next, note that $F(x) = \inf\{\sup\{f_k(x) : k \geq n\} : n \in \mathbb{N}\}$ and $G(x) = \sup\{\inf\{f_k(x) : k \geq n\} : n \in \mathbb{N}\}$, hence they are measurable by the first two theorems.

□

Remark. It is also true that for a measurable function $f : S \to \overline{\mathbb{R}}$ is measurable implies

$$f^+(x) = \sup\{f(x), 0\}$$
$$f^-(x) = \sup\{-f(x), 0\}$$

are also measurable.

9 Simple Functions

Definition 9.1. Let $S \subseteq \mathbb{R}$. Then,

$$\chi S : \mathbb{R} \longrightarrow \mathbb{R}$$
$$x \longmapsto \chi_S(x) = \begin{cases} 1, & x \in S \\ 0, & x \notin S \end{cases}$$

is the **characteristic function of** S .
A measurable function $s : \mathbb{R} \to \overline{\mathbb{R}}$ is a **simple functions** if $s(\mathbb{R})$ is finite.

Proposition 9.1. If s is a simple function. Then, there exists a finite, disjoint collection of measurable sets $\{S_k : 1 \leq k \leq K\}$ and a finite sequence of distinct real numbers $(a_k)_{1 \leq k \leq K}$ such that $\mathbb{R} = \bigcup_{k=1}^{K} S_k$ and $s = \sum_{k=1}^{K} a_k \chi_{S_k}$. Furthermore, this combination is unique up to permutation of the a_k, s_k. This representation is called the **canonical representation**.

Lemma 9.1. Let $f : \mathbb{R} \to \mathbb{R}$ be nonnegative and measurable with $f(\mathbb{R})$ being bounded, then for each $\varepsilon > 0$ there is a nonegative simple function s such that $f \geq s$ and $f(x) - s(x) < \varepsilon$ for all $x \in \mathbb{R}$.

Proof. There is a $M > 0$ such that $f(\mathbb{R}) \subseteq [0, M)$. Given ε, let $y_k = k\varepsilon$ for $k \in \mathbb{N}_0$. Since, $y_k - y_{k-1} = \varepsilon$, there is $N \in \mathbb{N}$ such that $[0, M] \subseteq \bigcup_{k \in \mathbb{N}} [y_{k-1}, y_k)$. Let $S_k = f^{-1}([y_{k-1}, y_k))$ for $1 \leq k \leq N$. Define $s = \sum_{k=1}^{N} y_{k-1} \chi_{S_k}$. Then, $s \geq 0$ and s is simple. Furthermore, for each $x \in \mathbb{R}$, there is a unique k, with $1 \leq k \leq N$ such that $f(x) \in [y_{k-1}, y_k)$. Consequently, $s(x) = y_{k-1} \leq f(x) < y_k$. Hence, $f(x) - s(x) < y_k - y_{k-1} = \varepsilon$. □

Theorem 9.1. $f : \mathbb{R} \to \overline{\mathbb{R}}$ is measurable if and only if there is a sequence of simple functions (s_n)a such that (s_n) converges pointwise to f and $|f| \geq |s_n|$ for all $n \in \mathbb{N}$.

Proof. Suppose the sequence (s_n). Then, f is measurable as

$$f = \lim_{n \to \infty} s_n = \limsup_{n \to \infty} s_n = \liminf_{n \to \infty} s_n.$$

Now, assume f is measurable. Then, $f = f^+ - f^-$. Both f^+ and f^- are measurable and nonnegative. Since the difference of two simple functions is simples, it suffices to assume $f \geq 0$, that is $f^- = 0$. Let $B_n = \{x \in \mathbb{R} : f(x) \leq n\}$ and $g_n = f\chi_{B_n}$ for all $n \in \mathbb{N}$. Since $g_n(x) = \inf\{f(x), n\chi_{B_n}\}$. Then, we see g_n is measurable as f and the simple function $n\chi_{B_n}$ are measurable. Furthermore, g_n is bounded. Hence, there is a measurable simple function r_n such that $g_n \geq r_n$ and $g_n(x) - r_n(x) < \frac{1}{n}$ for all x. Finally, define

$$s_n = r_n + n\chi_{B_n^c}.$$

Then, we find (s_n) is the sequence of functions desired. □

Corollary 9. Let (f_n) be a sequence of nonnegative measurable functions $f_n : \mathbb{R} \to \overline{\mathbb{R}}$. Then, $x \mapsto \sum_{i=1}^{\infty} f_k(x)$ is measurable. In particular, if $f, g : \mathbb{R} \to \overline{\mathbb{R}}$ are nonnegative and measurable, then so is $f + g$.

Proof. For $N \in \mathbb{N}$, let $F_n = \sum_{k=1}^{N} f_k$. For each k there is sequence of simple functions $(s_{k,n})_n$ such that $(s_{k,n})_n$ converges pointwise to f_k and $f_k \geq s_{k,n} \geq 0$

for all n.
Hence, $\left(\sum_{k=1}^{N} s_{k,n}\right)_n$ is a sequence of nonnegative simple functions such that $F_N \leq \sum_{k=1}^{N} s_{k,n}$ for all n and

$$\lim_{n\to\infty} \sum_{k=1}^{N} s_{k,m}(x) = F_N(x)$$

for all $x \in \mathbb{R}$.
So, F_N is the limit of a sequence of measurable functions, so it is measurable.
Furthermore, we have that for each $x \in \mathbb{R}$, $\left(F_{N(x)}\right)_N$ is increasing, we find

$$\sum_{k=1}^{\infty} f_k = \limsup_{N\to\infty} F_N = \lim_{N\to\infty} F_N.$$

□

10 Littlewood's 3 Principles

Remark.
1. Every measurable set is "nearly" the union of a finite collection of intervals.
2. Every measurable function is "nearly" continuous.
3. Every pointwise convergent sequence of measurable functions is "nearly" uniformly continuous.

We state these princeiples rigorously in the following way:

Theorem 10.1. If S is measurable, with $\mu(S) < \infty$, then for each $\varepsilon > 0$ there is a finite disjoint collection of open intervals $\{I_k : 1 \leq k \leq n\}$ such that for $U = \bigcup_{k=1}^{n} I_k$ we find

$$\mu(S \triangle U) < \varepsilon.$$

Theorem 10.2 (Lucin's Theorem)**.** Let $f : S \to \mathbb{R}$ be measurable with $\mu(S) < \infty$. Then, for each $\varepsilon > 0$ there is a compact $K \subseteq S$ such that $f\mid_K : K \to \mathbb{R}$ is continuous and $\mu(S \setminus K) < \varepsilon$.

Theorem 10.3 (Lucin's Theorem for functions on $\mathbb{R}$)**.** Let $f : \mathbb{R} \to \mathbb{R}$ be measurable. Then, for all $\varepsilon > 0$ there is a continuous $g : \mathbb{R} \to \mathbb{R}$ and a closed set $E \subseteq \mathbb{R}$ such that $f = g$ on E and $\mu(E^c) < \varepsilon$. Moreover, $\sup\{|g(x)| : x \in \mathbb{R}\} \leq \sup\{|f(x)| : x \in \mathbb{R}\}$.

Theorem 10.4 (Egoroff's Theorem)**.** Let S be measurable with $\mu(S) < \infty$. Suppose (f_n) is a sequence of measurable functions $f_n : S \to \mathbb{R}$ which converges pointwise almost everywhere to $f : S \to \mathbb{R}$. Then, for all $\varepsilon > 0$, there is a measurable $E \subseteq S$ such that $\mu(E) < \varepsilon$ and (f_n) converges uniformly to f on $S \setminus E$.

Lecture 16: Conclusion of Measure Theory and Lebesgue Integration

Tue 19 Oct 2021 13:02

Recall. We stated the theorems behind littlewood's 3 principles, now we prove them.

Proof.

1. (2.2). Let J be the collection of all open intervals (a, b) with $a, b \in \mathbb{Q}$ and $a < b$. Since J is countable we can order the intervals $J = \{J_k : k \in \mathbb{N}\}$. Let $\varepsilon > 0$ and first we do the case S is bounded. For each $n \in \mathbb{N}$, there is a closed set $C_n \subseteq f^{-1}(J_n)$ and a $D_n = S \setminus f^{-1}(I_n)$ such that $\mu(S \setminus (C_n \cup D_n)) < \frac{\varepsilon}{2^n}$. Since S is bounded, C_n and D_n are compact. Let $K = \bigcap_{n \in \mathbb{N}} (C_n \cup D_n)$ and as $C_n, D_n \subseteq S$, we see $K \subseteq S$. Furthermore, K is compact and we find $\mu(S \setminus K) \leq \sum_{i=1}^{\infty} \mu(S \setminus (C_n \cup D_n)) < \varepsilon$. Now, we show the restriction is continuous. Let $\varepsilon > 0$, then for $x \in K$ we find $a, b \in \mathbb{Q}$ such that $a < f(x) < b$ and $b - a < \varepsilon$. Hence, there is $n \in \mathbb{N}$ such that $I_n = (a, b)$. Consequently, $x \in f^{-1}(I_n)$ and $x \notin S \setminus f^{-1}(I_n)$. So, $x \in \left(S \setminus f^{-1}(I_n)\right)^c \subseteq D_n^c$. As D_n is closed, D_n^c is open, hence there is a $\delta > 0$ so that $(x - \delta, x + \delta) \subseteq D_n^c$. If $y \in K \cap D_n^c$, then $y \in C_n$, thus $y \in f^{-1}(I_n)$, hence $a < f(y) < b$. So, $|f(x) - f(y)| < b - a = \varepsilon$ for $y \in (x - \delta, x + \delta)$.
   Now, we do the unbounded case. As S is unbounded and $\varepsilon > 0$, we find $N \in N$ so that $S' = S \cap [-N, N]$ has the property $\mu(S \setminus S') < \frac{\varepsilon}{2}$, that is S is approximated by a bounded function arbitrarily well. Since S' is bounded, there is a compact set $K \subseteq S' \subset S$ so that $f \mid K$ is continuous and $\mu(S' \setminus K) < \frac{\varepsilon}{2}$. Then, $\mu(S \setminus K) = \mu(S \setminus S') + \mu(S' \setminus K) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

2. (2.4). Let E^* be the set of all $x \in S$ such that $(f_n(x))$ does not converge. By assumption, $\mu(E^*) = 0$. Since $f(x) = \lim\limits_{n \to \infty} f_n(x) = \limsup_{n \to \infty} f_n(x)$ for all $x \in S \setminus E^*$, then f is measurable. For $k, \ell \in \mathbb{N}$, let $E_{k,\ell} = \{x \in S : |f_\ell(x) - f(x)| \geq \frac{1}{k}\}$. Then, $E_{k,\ell}$ is measurable. Fix k. If for each $n \in \mathbb{N}$ there is a $\ell \geq n$ so that $|f_\ell(x) - f(x)| \geq \frac{1}{k}$, then $x \in E^*$ as f does not converge at that point. Hence, $\bigcap_{n \in \mathbb{N}} \bigcup_{\ell=n}^{\infty} E_{k,\ell} \subseteq E^*$. Since $\mu\left(\bigcup_{\ell=1}^{\infty} E_{k,\ell}\right) \leq \mu(S) \leq \infty$, and the collection $\{\bigcup_{\ell=n}^{\infty} E_{k,\ell}\}$ is clearly descending. Hence, $\mu\left(\bigcap_{n \in \mathbb{N}} \bigcup_{k=n}^{\infty} E_{k,\ell}\right) = \lim\limits_{n \to \infty} \mu\left(\bigcup_{\ell=n}^{\infty} E_{k,l}\right) \leq \mu(E^*) = 0$. This holds for all $k \in \mathbb{N}$. So, for $\varepsilon > 0$ and $k \in \mathbb{N}$, we have a $n_k \in \mathbb{N}$ such that $\mu\left(\bigcup_{\ell=n_k}^{\infty} E_{k,\ell}\right) < \frac{\varepsilon}{2^k}$. Thus, $E = \bigcup_{k \in \mathbb{N}} \bigcup_{\ell=n_k}^{\infty} E_{k,\ell}$ is measurable and $\mu(E) < \sum_{k=1}^{\infty} \bigcup_{\ell=n_k}^{\infty} E_{k,l} = \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon$. If $x \in S \setminus E$, then $|f_n(x) - f(x)| < \frac{1}{k}$ for $k \in \mathbb{N}$ if $n \geq n_k$. So, (f_n) converges uniformly on $S \setminus E$.

□

This concludes measure theory.

11 Lebesque Integration

Definition 11.1 (Lebesque Integral: Nonnegative Simple Functions)**.** Let s be a nonnegative simple function of the form $s = \sum_{k=1}^{K} a_k \chi_{S_k}$ where $\{S_k : 1 \leq k \leq K\}$ is a disjoint collection of measurable sets. Then, the **Lebesque Integral** of s is defined to be

$$\int s = \int s(x)\,dx = \int s d\mu = \sum_{k=1}^{K} a_k \mu(S_k).$$

Proposition 11.1. If s is nonnegative and simple with two representations, $s = \sum_{k=1}^{K} a_k \chi_{S_k} = \sum_{j=1}^{J} b_j \chi_{T_j}$ for disjoint collections of measurable sets $\{S_k : 1 \leq k \leq K\}$ and $\{T_j : 1 \leq j \leq J\}$. Then

$$\sum_{k=1}^{K} a_k \mu(S_k) = \sum_{j=1}^{J} b_j \mu(T_j).$$

In particular, $\int s$ is well defined.

The proof of this is trivial.

Lemma 11.1. Let s, t be nonnegative and simple and $\alpha \geq 0$. Then

$$\alpha \cdot \int s = \int \alpha \cdot s \text{ and } \int (s+t) = \int s + \int t$$

.

Proof. Clearly, multiplying the sum times α yields $\alpha \sum_{k=1}^{K} a_k \mu(S_k) = \sum_{k=1}^{K} \alpha a_k \mu(S_k)$. For the second claim. Suppose $s = \sum_{k=1}^{K} a_k \chi_{S_k}$ and $g = \sum_{j=1}^{J} b_j \chi_{T_j}$ are canonical representations. Then, $s + t = \sum_{k=1}^{K} \sum_{j=1}^{J} (a_k + b_j) \chi_{S_k \cap T_j}$ with $\{S_k \cap T_j : 1 \leq k \leq K, 1 \leq j \leq J\}$ is a disjoint collection and

$$\begin{aligned}
\int (s+t) &= \sum_{k=1}^{K} \sum_{j=1}^{J} (a_k + b_j) \mu(S_k \cap T_j) \\
&= \sum_{k=1}^{K} a_k \sum_{j=1}^{J} \mu(S_k \cap T_j) + \sum_{j=1}^{J} b_j \sum_{k=1}^{K} \mu(S_k \cap T_j) \\
&= \sum_{k=1}^{K} a_k \mu(S_k) + \sum_{j=1}^{J} b_j \mu(T_j) \\
&= \int s + \int t.
\end{aligned}$$

□

Lemma 11.2. Let s, t be nonnegative and simple such that $s \leq t$. Then, $\int s \leq \int t$.

Proof.

$$\begin{aligned} \int t &= \int (t - s + s) \\ &= \int \underbrace{(t-s)}_{\geq 0} + \int s \\ &\geq \int s. \end{aligned}$$

□

Definition 11.2. Let $f : S \to \overline{\mathbb{R}}$, then the **zero extension** of f to $\mathbb{R}$ is

$$\begin{aligned} f^* : \mathbb{R} &\longrightarrow \overline{\mathbb{R}} \\ x &\longmapsto f^*(x) = \begin{cases} f(x), & x \in S \\ 0, & x \notin S \end{cases}. \end{aligned}$$

Moreover, this function preserves measurability.

Definition 11.3 (Lebesque Integral of a General Nonnegative Function)**.** Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be a nonnegative measurable function and $\mathscr{S}(f)$ be the collection of all nonnegative simple functions, s, such that $s \leq f$. Then, the **Lebesque Integral** of f over $\mathbb{R}$ is defined to be

$$\int f = \int_{\mathbb{R}} f(x)\, dx = \sup\{\int s : s \in \mathscr{S}(f)\}$$

.

If $f : S \to \overline{\mathbb{R}}$ is nonnegative and measurable, then

$$\int_S f = \int_S f(x)\, dx = \int_{\mathbb{R}} f^*$$

.

Theorem 11.1 (Chebyshev's Inequality)**.** Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be nonnegative and measurable. Then, for any $\lambda \in (0, \infty)$, then

$$\mu(\{x \in \mathbb{R} : f(x) \geq \lambda\}) \leq \frac{1}{\lambda} \int f.$$

Proof. Let $E = \{x \in \mathbb{R} : f(x) \geq \lambda\}$. This is the preimage of an extended borel set, hence measurable. Let $s = \lambda \chi_E$. Then, $s \in \mathscr{S}(f)$. Hence, $\int s = \lambda \mu(E) \leq \int f$. Hence the inequality holds. □

Theorem 11.2. Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be nonnegative ad measurable. Then $\int f = 0$ if and only if $f(x) = 0$ for almost every $x \in \mathbb{R}$.

Proof. Suppose $\int f = 0$, then by chebyshev

$$\begin{aligned}\mu(\{x \in \mathbb{R} : f(x) > 0\}) &= \mu(\bigcup_{n\in\mathbb{N}} \{x \in \mathbb{R} : f(x \geq n)\}) \\ &\leq \sum_{n=1}^{\infty} \mu\left(\{x \in \mathbb{R} : f(x) \geq \frac{1}{n}\}\right) \\ &= \sum_{i=1}^{\infty} \int f \\ &= 0.\end{aligned}$$

Conversely, if $f(x) = 0$ almost everywhere, then for every $s \in \mathscr{S}(f)$, we see s is zero almost everywhere, hence $\int s = 0$, so $\int f = \sup\{0 : s \in \mathscr{S}(f)\} = 0$. ☐

Lecture 17: General Lebesque Integral

Tue 26 Oct 2021 13:15

Lecture 18: General Lebesque Integral (2)

Tue 26 Oct 2021 13:16

Proposition 11.2. Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be integrable. Then for each $\varepsilon > 0$ there is a $\delta > 0$ such that each measurable $S \subseteq \mathbb{R}$ has $\int_S |f| < \varepsilon$ if $m(S < \delta)$.

Proof. Let $\varepsilon > 0$, then there is a $s \in \mathscr{S}(|f|)$ such that $\int (|f| - s) < \frac{\varepsilon}{2}$. Let $a\alpha = \sup\{s(x) : x \in \mathbb{R}\}$ and $\delta = \frac{\varepsilon}{2(\alpha+\varepsilon)}$. If S is measurable and $m(S) < \delta$, we find

$$\int_S |f| \leq \int s + \frac{\varepsilon}{2} \leq \alpha m(S) + \frac{\varepsilon}{2} < \varepsilon.$$

☐

Theorem 11.3 (Monotone Convergence Theorem)**.** Let (f_n) be a sequence of nonnegative measurable functions with $f_n : \mathbb{R} \to \overline{\mathbb{R}}$ such that $(f_n(x))$ is increasing for all $x \in \mathbb{R}$. Then, $f = \lim_{n\to\infty} f_n$ is maesurable with $\int f = \lim_{n\to\infty} \int f_n$.

Proof. Since $f = \limsup_{n\to\infty} f_n = \liminf_{n\to\infty} f_n$, we see f is measurable. Moreover, the sequence $\left(\int f_n\right)$ is increasing (as the f_ns are increasing). Hence, letting $L = \lim_{n\to\infty} \int f_n$ exists with $L \in R_0^+$. Since $\int f_n \leq \int f$ for all n by monotonicity, we find $L \leq \int f$.
Let $s \in \mathscr{S}(f)$ and fix $c \in (0, 1)$ and define $E_n = \{x \in \mathbb{R} : f_n(x) \geq cs(x)\}$. Then, we find $\{E_n : n \in \mathbb{N}\}$ is an ascending collection (again by monotonicity

of (f_n)) of measurable sets with $\bigcup_{n\in\mathbb{N}} E_n = \mathbb{R}$ as $cs(x) < f_n(x) \leq f(x)$. Let $s = \sum_{k=1}^{K} a_k \chi_{S_k}$ and we see $cs\chi_{E_n} M = f_n \chi_{E_n} \leq f_n$, with

$$L \geq \int f_n \geq \int_{E_n} f_n \geq \int cs\chi_{E_n} = c\int_{E_n} s = c\sum_{k=1}^{K} a_k m(S_k \cap E_n).$$

Since $\lim_{n\to\infty} m(E_n \cap S_n) = m(S)$ for every measurable set S, we find $L \geq c\sum_{k=1}^{K} a_k m(S_k) = c\int s$. Since c was arbitrary, we see the inequality holds for all $c \in (0,1)$, hence we find $L \geq s$ (by taking supremums), but $s \in \mathscr{S}(f)$, hence $L \geq \int f$. So, $L = \int f$. □

Theorem 11.4 (Fatou's Lemma)**.** If (f_n) is a sequence of nonnegative measurable functions $f_n : \mathbb{R} \to \overline{\mathbb{R}}$, then $\int \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int f_n$.

Proof. For $x \in \mathbb{R}$, define $g_n(x) = \inf\{f_k(x) : k \geq n\}$ for $n \in \mathbb{N}$. Then, we find (g_n) is a nonnegative measurable sequence of functions with $(g_n(x))$ increasing for all fixed x and $g_n \leq f_n$ for all n. Consequently, $\int g_n \leq \int f_n$ and $\left(\int g_n\right)$ is increasing. As $\lim_{n\to\infty} g_n = \liminf_{n\to\infty} f_n$ is measurable by an earlier theorem, we find

$$\liminf_{n\to\infty} \int f_n \geq \liminf_{n\to\infty} \int g_n = \lim_{n\to\infty} \int g_n = \int \lim_{n\to\infty} g_n = \int \liminf_{n\to\infty} f_n.$$

□

Proposition 11.3. For any integral function $f : \mathbb{R} \to \overline{\mathbb{R}}$, we find $\left|\int f\right| \leq \int |f|$.

Theorem 11.5 (Dominated Convergence Theorem)**.** Let (f_n) be a sequence of measurable functions $f_n : \mathbb{R} \to \overline{\mathbb{R}}$. Suppose there is an integrable function g with $|f_n| \leq g$ for all $n \in \mathbb{N}$. If (f_n) converges pointwise to a function $f : \mathbb{R} \to \overline{\mathbb{R}}$ almost everywhere, then f is integrable and

$$\lim_{n\to\infty} \int |f_n - f| = 0 \text{ and } \lim_{n\to\infty} \int f_n = \int f.$$

Proof. Since $f(x) = \lim_{n\to\infty} f_n(x)$ for almost all $x \in R$, we find f is measurable. Moreover, $|f_n| \leq g$ implies $|f| < g$ almost everywhere and since g is integrable (hence finite a.e) we find f, f_n are integrable (hence finite) almost everywhere. Now, define for each $n \in \mathbb{N}$

$$E_n = \{x \in \mathbb{R} : |f_n(x)|, |f(x)| < \infty, |f_n(x) - f(x)| \leq 2g(x)\}.$$

Since $R \setminus \bigcup_{n\in\mathbb{N}} E_n$ is a set of measure 0, we can assume $|f_n(x)|, |f(x)| < \infty$ and $|f_n() - f(x)| \leq 2g(x)$ for all $x \in \mathbb{R}$. Then, Fatou's lemma applies to the sequence on nonnegative measurable functions $(2g - |f_n - f|)$ yielding

$$\begin{aligned}\int 2g &\leq \liminf_{n\to\infty}(2g - |f_n - f|)\\ &= \int 2g + \liminf_{n\to\infty}\left(-\int |f_n - f|\right)\\ &= \int 2g - \limsup_{n\to\infty}\int |f_n - f|\\ &\Rightarrow \limsup_{n\to\infty}\int |f_n - f| \leq 0\\ &\Rightarrow \lim_{n\to\infty}\int |f_n - f| = 0.\end{aligned}$$

Hence, $\lim\limits_{n\to\infty}\left|\int (f_n - f)\right| = 0$ by the earlier lemma. So, $\lim\limits_{n\to\infty}\int f_n = \int f_n$. □

Definition 11.4 (Convergence in Measure)**.** Let (f_n) be a sequence of measurable functions $f_n : \mathbb{R} \to \overline{\mathbb{R}}$ and $f : \mathbb{R} \to \overline{\mathbb{R}}$ also be measurable. The sequence (f_n) **converges in measure** to f ($f_n \to f$ by measure) if each $f_n$is finite almost everywhere and for each $\varepsilon > 0$ there is a $N \in \mathbb{N}$ so that

$$m\left(\{x \in \mathbb{R} : |f_n(x) - f(x)| > \varepsilon\}\right) < \varepsilon$$

for $n \geq N$.

Theorem 11.6 (Riesz)**.** Let (f_n) be a sequence of measurable functions $f_n : \mathbb{R} \to \overline{\mathbb{R}}$ and $f : \mathbb{R} \to \overline{\mathbb{R}}$ also being measurable. If $(f_n) \to f$ in measure, then there is a subsequence (f_{n_k}) which converges pointwise almost everywhere to f.

Proof. First, we find a strictly increasing sequence of numbers (n_k) such that $m\left(\{x \in \mathbb{R} : |f_j(x) - f(x)| > 2^{-k}\}\right) < 2^{-k}$ if $j \geq n_k$. For $k \in \mathbb{N}$ denote

$$S_k = \{x \in \mathbb{R} : |f_{n_k} - f(x)| > 2^{-k}\}.$$

Then, $\sum_{k=1}^{\infty} m(S_k) \leq \sum_{k=1}^{\infty} 2^{-k} < \infty$. Applying the Borel-Cantelli Lemma yields that almost every $x \in R$ does not belong to any infinite subcollections of (S_k). For such x, we find a $K \in \mathbb{N}$ such that $|f_{n_k}(x) - f(x)| \leq 2^{-k}$ for $k \geq K$. Hence, f_{n_k} converges pointwise to f for all x not belonging to an infinite subcollection of (S_k), hence almost everywhere. □

Lecture 19: End of Convergence, Functions of Bounded Variation, and Derivatives

Thu 28 Oct 2021 13:02

Recall we had the dominated convergence theorem. A similair version of the theorem makes use of convergence in measure as follows.

Theorem 11.7 (Dominated Convergence - Convergence in Measure)**.** Let (f_n) be a sequence of measurable functions $f_n : \mathbb{R} \to \overline{\mathbb{R}}$ and suppose there is an integrable function $g : \mathbb{R} \to \overline{\mathbb{R}}$ so that $|f_n| \leq g$ for all $n \in \mathbb{N}$. If $(f_n) \to f : \mathbb{R} \to \overline{\mathbb{R}}$ in measure, (with f measurable), then f is integrable and $\lim\limits_{n\to\infty} \int |f_n - f| = 0$ and $\lim\limits_{n\to\infty} \int f_n = f$.

Proof. First, note a subsequence of (f_n) converges to f pointwise almost everywhere. Hence, we find $|f| \leq g$ almost everywhere, so f is integrable. We cam assume $|f_n - f| \leq 2g$ (almost) everywhere. Then, we find a subsequence $(g_n) = (f_{n_k})$ such that $\limsup_{n\to\infty} |f_n - f| = \lim\limits_{n\to\infty} |g_k - f|$. Then, as $(g_k) \to f$ in measure, we find another subsequence $(h_j) = \left(g_{k_j}\right) = \left(f_{n_{k_j}}\right)$ which converges pointwise to f almost everywhere.
Applying dominated convergence theorem yields

$$\lim_{n\to\infty} \int |h_j - f| = 0.$$

Then, we find

$$\begin{aligned}\limsup_{n\to\infty} \int |f_n - f| &= \lim_{n\to\infty} \int |g_k - f| \\ &= \lim_{n\to\infty} |h_j - f| \\ &= 0.\end{aligned}$$

This completes the proof. □

12 Functions of Bounded Variation and Absolutely Continuous Functions

Remark. For this chapter $[a, b] \subseteq R$ will always denote a compact interval on $\mathbb{R}$.

Definition 12.1 (Partition)**.** A finite sequence $P = (x_k)_{k=n}^{N}$ with $n, N \in \mathbb{Z}$ and $n < N$ is called a **partition** of $[a, b]$ if $x_n = a$, $x_N = b$ and $x_{k-1} \leq x_k$ for $n < k \leq N$. We denote the collection of all partitions of $[a, b]$ to be $\mathscr{P}([a, b])$.

Definition 12.2. Let $f : [a, b] \to \mathbb{R}$ be a function. Then,

- For a partition $P = (x_k)_{k=n}^{N}$, we denote

$$V(f, P) = \sum_{k=n+1}^{N} \left| f(x_k) - f_{(x_{k-1})} \right|$$

  to be the **variation of f with respect to P**.
- We define the quantity $\mathrm{TV}(f) = \sup\{V(f, P) : P \in \mathscr{P}([a, b])\}$ to be the **total variation of f** .

Remark. If $f : [a, b] \to \mathbb{R}$ and $c \in [a, b]$ with partitions $P_1 = (x_k)_{k=n}^{N}$ of $[a, c]$ and $P_2 = (x_k)_{k=N}^{K}$ of $[c, b]$. Then denote, $P = (x_k)_{k=n}^{K}$ to be a partition of $[a, b]$ and we find

$$V(f, P) = V\left(f \mid_{[a,c], P_1}\right) + V\left(f \mid_{[c,b]}, P_2\right).$$

Moreover,

$$\mathrm{TV}(f) = \mathrm{TV}\left(f \mid_{[a,c]}\right) + \mathrm{TV}\left(f \mid_{[c,b]}\right).$$

Definition 12.3 (Bounded Variation)**.** A function $f : \mathbb{R} \to \overline{\mathbb{R}}$ has **bounded variation** if $\mathrm{TV}(f) < \infty$.

Theorem 12.1 (Jordan's Theorem)**.** A function $f : [a, b] \to \mathbb{R}$ is of bounded variation if and only if there are increasing functions $g, h : [a, b] \to \mathbb{R}$ so that $f = g - h$.

Proof. Suppose $\mathrm{TV}(f) < \infty$ and let $x, y \in [a, b]$ with $x < y$. Then, we find

$$\begin{aligned} \mathrm{TV}\left(f \mid_{[a,y]}\right) &= \mathrm{TV}\left(f \mid_{[a,x]}\right) + \mathrm{TV}\left(f \mid_{[x,y]}\right) \\ &\geq \mathrm{TV}\left(f \mid_{[a,x]}\right) + |f(y) - f(x)| \\ &\geq \mathrm{TV}\left(f \mid_{[a,x]}\right) + f(x) - f(y). \end{aligned}$$

Furtheromre, $h : x \mapsto \mathrm{TV}\left(f \mid_{[a,x]}\right)$ and $g : x \mapsto \mathrm{TV}\left(f \mid_{[a,x]}\right) + f(x)$ are increasing. This fact is trivial for h and we find , adding $f(y)$ to both sides of the former inequality yields $g(y) \geq g(x)$ for arbitrary $y \geq x$, so this claim holds as well.

Taking the difference, $g - h = f$.

Conversely, suppose $f = g - h$ for increasing $g, h : [a, b] \to \mathbb{R}$. Then, let $x, y \in [a, b]$ with $y \geq x$. Then, we find

$$\begin{aligned} |f(y) - f(x)| &= |g(y) - g(x) + h(x) - h(y)| \\ &\leq |g(y) - g(x)| + |h(x) - h(y)| \\ &= g(y) - g(x) + h(y) - h(x). \end{aligned}$$

Hence, for a partition $P = (x_k)_{k=n}^{N}$, we find

$$\begin{aligned} V(f,P) &= \sum_{k=n+1}^{N} |f(x_k) - f(x_{k-1})| \\ &\le \sum_{k=n+1}^{N} (g(x_k) - g(x_{k-1}) + h(x_k) - h(x_{k-1})) \quad = g(b) - g(a) + h(b) - h(a) \\ &< \infty. \end{aligned}$$

□

Definition 12.4 (Absolute Continuity)**.** A function $f : [a,b] \to \mathbb{R}$ is **absolutely continuous** if for each $\varepsilon > 0$ we find a $\delta > 0$ such that for every finite disjoint collection of nonempty intervals $\{(a_k, b_k) \subseteq [a,b] : 1 \le k \le K\}$ with $\sum_{k=1}^{K} (b_k - a_k) < \delta$, we have $\sum_{k=1}^{K} |f(a_k) - f(b_k)| < \varepsilon$.

Remark. Absolute continuity is stronger than uniform continuity, but weaker than lipschitz continuity.

Theorem 12.2. If a function $f : [a,b] \to \mathbb{R} \to$ is absolutely continuous, then f is continuous and f has bounded variation.

Proof. f is trivially continuous, taking a finite disjoint collection consisting only of 1 interval $\{(x,y)\}$ yields the definition of continuity.
Now we show bounded variation. For $\varepsilon = 1$, let $\delta > 0$ be the number such that the definition of absolute continuity holds for f.
Now fix $(x_k)_{k=n}^{N} \in \mathscr{P}([a,b])$ so that $x_k - x_{k-1} < \delta$ for all $n < k \le N$. Then, if $P \in \mathscr{P}([x_{k-1}, x_k])$, we see $V\left(f|_{[x_{k-1},x_k]}, P\right) < 1$ by definition of absolute continuity.
So, we have $\mathrm{TV}([x_{k-1}, x_k]) \le 1$, so $\mathrm{TV}(f) = \sum_{k=n+1}^{N} \mathrm{TV}\left(f|_{[x_{k-1},x_k]}\right) \le N - n$ by the ε assumption. □

As it turns out, absolutely continuous functions have a relation to integrable functions, particularly, an integrable function f is simply the anti-integral of an absolutely continuous one.

Proposition 12.1. If $f : [a,b] \to \overline{\mathbb{R}}$ is integrable, then,

$$F : [a,b] \to \mathbb{R}, \; x \mapsto \int_{[a,x]} f$$

is absolutely continuous.

This claim can be generalized into a sort of fundamental theorem of calculus for the lebesque integrals to characterize integrals and derivatives. For now, we only prove the weak version.

Proof. For $\varepsilon > 0$ there is a $\delta > 0$ such that $\int_S |f| < \varepsilon$ for every measurable set $S \subseteq [a,b]$ with $m(S) < \delta$.
Now, let $\{(a_k, b_k) : 1 \leq k \leq K\}$ be a disjoint collection of intervals in $[a,b]$ with $\sum_{k=1}^{K} (b_k - a_k) < \delta$. Fix $S = \bigcup_{k=1}^{K} (a_k, b_k)$. Then, since $m(S) < \delta$ and

$$\begin{aligned}
\sum_{k=1}^{K} |F(b_k) - F(a_k)| &= \sum_{k=1}^{K} \left| \int_{[a_k, b_k]} f \right| \\
&\leq \sum_{k=1}^{K} \int_{[a_k, b_k]} |f| \\
&= \int_S |f| \\
&< \varepsilon \text{ by assumption.}
\end{aligned}$$

Hence, absolute continuity holds. □

13 Derivatives and Fundamental Theorem of Calculus

Proposition 13.1. Let $f : (a,b) \to \overline{\mathbb{R}}$ be monotone on $(a,b) \subseteq \mathbb{R}$ with $a, b \in \overline{\mathbb{R}}$ and $a < b$. Then,

$$\lim_{x \to a} f(x) = \inf\{f(x) : x \in (a,b)\}, \lim_{x \to b} f(x) = \sup\{f(x) : x \in (a,b)\}$$

are both well defined.

Lecture 20: Derivatives (2)

Tue 02 Nov 2021 12:58

Recall. A monotone function on an interval has well defined limits at both its endpoints.

Definition 13.1 (Upper/Lower Derivatives)**.** Let $S \subseteq \mathbb{R}$, $f : S \to \mathbb{R}$

- We define $\overline{D}f(x) = \lim_{\tau \to 0} \sup\{\frac{f(x+h)-f(x)}{h} : 0 < |h| < \tau\}$ to be the **upper derivative**.
- We define $\underline{D}f(x) = \lim_{\tau \to 0} \inf\{\frac{f(x+h)-f(x)}{h} : 0 < |h| < \tau\}$ to be the **lower derivative**.
- If, for some $x \in \overset{\circ}{S}$, we find $\overline{D}f(x), \underline{D}f(x) \in \mathbb{R}$, with the upper and lower derivatives being equal, we say f is **differentiable** at x.
  We denote $f'(x) = \overline{D}f(x) = \underline{D}f(x)$.

We know, the limits of the upper and lower derivatives to be well defined as the supremum and infimum are monotone functions with respect to τ.

Proposition 13.2. Let $f : S \to \mathbb{R}$ and let $x \in \overset{\circ}{S}$. Then, f is differentiable at x if and only if

$$\lim_{y\to x} \frac{f(y) - f(x)}{y - x} = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \in \mathbb{R}.$$

That is, the classical derivative is equivalent to the lebesque derivative, so we will use the new definition for most proofs, but the old for most computations.

Theorem 13.1 (Mean-Value Theorem)**.** Let $f : [a,b] \to \mathbb{R}$ be continuous and differentiable at every $x \in (a,b)$. Then, there exists $\xi \in (a,b)$ so that $f(b) - f(a) = f'(\xi)(b-a)$.

Lemma 13.1. Let $f : [a,b] \to \mathbb{R}$ be increasing and suppose $\overline{D}f(x) = \underline{D}f(x)$ for almost every $x \in [a,b]$. Then, $\overline{D}f(x)$ and $\underline{D}f(x)$ are finite almost everywhere. Moreover, f is differentiable almost everywhere (on $[a,b]$). Furthermore, f' is an integrable function and

$$\int_{[a,b]} f' \leq f(b) - f(a).$$

Proof. Extend f to $[a,\infty)$ by letting $f(c) = f(b)$ for all $c \geq b$. Define a sequence (g_n), $g_n : [a,b] \to \overline{\mathbb{R}}$ with

$$x \mapsto n\left(f\left(x + \frac{1}{n}\right) - f(x)\right).$$

Then, b assumption, we know $(g_n(x))$ to be convergent in $\overline{\mathbb{R}}$ with limit $f'(x)$ for almost every $x \in (a,b)$. Each g_n is measurable, hence $\lim_{n\to\infty} g_n$ is increasing, we see $g(n) \geq 0$, hence $\overline{D}f \geq 0$.
Applying Fatou's lemma yields

$$\begin{aligned}
\int_{[a,b]} \overline{D}f &= \int_{[a,b]} \liminf_{n\to\infty} f_n \\
&\leq \liminf_{n\to\infty} \int_{[a,b]} g_n \\
&= \liminf_{n\to\infty} n\left(\int_{\left[a+\frac{1}{n},b+\frac{1}{n}\right]} f - \int_{[a,b]} f\right) \\
&= \liminf_{n\to\infty} \left(\underbrace{n\int_{\left[b,b+\frac{1}{n}\right]} f}_{=f(b)} - \underbrace{n\int_{\left[a,a+\frac{1}{n}\right]} f}_{\leq f(a)}\right) \\
&\leq f(b) - f(a).
\end{aligned}$$

We know the final inequality holds because f is constant on $\left[b, b+\frac{1}{n}\right]$ and though f is not constant, it is increasing on $\left[a, a+\frac{1}{n}\right]$ hence the upper bound of their difference is attained by $f(a)$.
Consequently, $\overline{D}f$ is integrable (so finite almost everywhere). And, since $\overline{D}f = \underline{D}f$, we find $f'(x)$ exists and equals $\overline{D}f(x)$ for almost every $x \in [a, b]$. $\square$

Later, we will prove equality holds precisely in the case of absolute continuity.

Definition 13.2 (Vitali Covering)**.** Let $S \subseteq \mathbb{R}$. We call a collection of closed, bounded intervals (denoted $\mathscr{C}$) of positive length a **Vitali covering** of $S \subseteq \mathbb{R}$ if for every $x \in S$ and $\varepsilon > 0$ we find an $I \in \mathscr{C}$ such that $x \in I$ and $l(I) < \varepsilon$.

Example. A vitali covering of $S = [0, 1]$ goes as follows. Let $H = \mathbb{Q} \cap [0, 1]$, then $\mathscr{C} = \{[x, x+h] : h \in H, x \in [0, 1]\}$. $\diamond$

Theorem 13.2 (Vitali Covering Lemma)**.** Let $\mathscr{C}$ be a Vitali covering of the set $S \subseteq \mathbb{R}$ with $m^*(S) < \infty$. Then, for every $\varepsilon > 0$ there is a finite, disjoint collection of intervals $\{I_k \in \mathscr{C} : 1 \leq k \leq n\}$ such that

$$m^*\left(S \setminus \bigcup_{k=1}^{n} I_k\right) < \varepsilon.$$

Theorem 13.3 (Lebesgue's Theorem)**.** Let $f : I \to \mathbb{R}$ be a monotone function on an interval $I \subseteq \mathbb{R}$. Then, f is differentiable at almost every $x \in I$ and f' is integrable on every interval $[a,b] \subseteq I$. In particular, if f is increasing, then
$$\int_{[a,b]} f' \leq f(b) - f(a).$$

Proof. It suffices to show I is open and bounded, else we could replace I by $\overset{\circ}{I} \cap (-n,n)$ for $n \in N$ and we find $\overset{\circ}{I} = \bigcup_{n\in\mathbb{N}} \overset{\circ}{I} \cap (-n,n)$. Similarly, we can assume f to be increasing. Hence, for all $x \in I$, we have $0 \leq \underline{D}f(x) \leq \overline{D}f(x) \leq \infty$. So, we need only show $\overline{D}f(x) = \underline{D}f(x)$ with this quantity being finite for almost every $x \in I$.
For $p, q \in \mathbb{Q}$ and $p > q > 0$, define $E_{p,q} = \{x \in I : \underline{D}f(x) < q < p < \overline{D}f(x) < \infty\}$. Then,
$$\{x \in I : \underline{D}f(x) < \overline{D}f(x) < \infty\} = \bigcup_{p,q\in Q^+} E_{p,q}.$$
If f fails to be differentiable at $x \in I$, then either $x \in E_{p,q}$ for some $p, q \in \mathbb{Q}$ or $\overline{D}f(x) = \infty$. We know $\overline{D}f$ to be finite almost everywhere, so by subadditivity, we need only show the other component, $E_{p,q}$, has measure 0.
Fix $p, q \in \mathbb{Q}$ and suppose $m^*(E_{p,q} = m_0)$. Then, $m_0 \in [0,\infty)$ by the boundedness assumption. Given $\varepsilon > 0$ there is a nonempty open U such that $E_{p,q} \subseteq U$ and $m(U) < m_0 + \varepsilon$. Suppose $x \in E_{p,q}$. Since $\underline{D}f(x) < q$ by definition of $E_{p,q}$; for every $\delta > 0$ we find a $0 < h < \delta$ such that $[x, x+h] \subseteq U$ and $f(x+h) - f(x) < qh$ or $[x-h, x] \subseteq U$ and $f(x) - f(x-h) \leq qh$.
The collection $\mathscr{L}$ of all such intervals $[x, x+h]$ or $[x-h, x]$ for a fixed $\delta > 0$ and $x \in E_{p,q}$ forms a Vitali covering of $E_{p,q}$. We find all intervals $[a,b] \in \mathscr{L}$ have the property $f(b) - f(a) < q(b-a)$ by the earlier observation. Then, by the Vitali covering lemma, there is a finite, disjoint collection of intervals $\{I_n \in \mathscr{L} : 1 \leq n \leq N\}$ such that for $V = \bigcup_{n=1}^N I_n$, we have $m^*(E_{p,q} \setminus V) < \varepsilon$. Note that $m(V) < m_0 + \varepsilon$ since $V \subseteq U$. Since $m^*(E_{p,q} \setminus V) + m^*(E_{p,q} \cap V) \geq m_0$ since the two sets together contain $E_{p,q}$, we have $m^*(E_{p,q} \cap V) \geq m_0 - \varepsilon$.
Now, we follow a similair construction. If $x \in E_{p,q} \cap V$, then $p < \overline{D}f(x)$ implies for all $\delta > 0$ there is an $0 < h < \delta$ such that $[x, x+h] \subseteq V$ and $f(x+h) - f(x) \geq ph$ or $[x-h, x] \subseteq V$ and $f(x) - f(x-h) \geq ph$. The collection $\mathscr{U}$ of all such intervals $[x, x+h]$ or $[x-h, x]$ for a fixed $\delta > 0$ and $x \in E_{p,q} \cap V$ is a vitali covering of $E_{p,q} \cap V$. Moreover, if $[c,d] \in \mathscr{U}$, then $f(d) - f(c) \geq p(d-c)$. Applying Vitali Covering lemma yields a finite disjoint collection of intervals $\{I_k \in \mathscr{U} : 1 \leq k \leq K\}$ such that for $W = \bigcup_{k=1}^K J_k$, we have $m^*((E_{p,q} \cap V) \setminus W) < \varepsilon$. Since
$$m^*((E_{p,q} \cap V) \setminus W) + m(W) \geq m^*(E_{p,q} \cap V)$$
we have that $m(W) \geq m_0 - 2\varepsilon$.
We know each interval $J_k = [c_k, d_k]$ from W must be contained in V, furthermore it is contained in an interval $I_n = [a_n, b_n]$ of V. As each interval is disjoint and monotonic, we must have that
$$\sum_{k=1}^{K} (f(d_k) - f(c_k)) \leq \sum_{n=1}^{N} (f(b_n) - f(a_n)).$$

Now, since $I_n \in \mathscr{L}$ and $J_k \in \mathscr{U}$, we have

$$\begin{aligned} p\sum_{k=1}^{K}(d_k - c_k) &= pm(w) \\ &\leq qm(V) \\ &= q\sum_{n=1}^{N}(b_n - a_n) \end{aligned}$$

.

Hence, $p(m_0 - 2\varepsilon) \leq q(m_0 + \varepsilon)$ for each $\varepsilon > 0$, so $pm_0 \leq qm_0$ and as $p > q$, we must have $m_0 = 0$, so f is differentiable on all but sets of measure 0, so it is differentiable almost everywhere. $\square$

Corollary 10. If the function $f : [a, b] \to \mathbb{R}$ is of bounded variation on the interval $[a, b] \subseteq \mathbb{R}$, then it is differentiable at almost every $x \in [a, b]$. Consequently, if f is absolutely continuous on $[a, b]$, then it is differentiable at almost every $x \in [a, b]$.

Proof. Bounded variation implies $f = g - h$ for increasing functions g, h. Applying lebesque's theorem yields g, h are differentiable almost everywhere, hence f is differentiable almost everywhere. $\square$

Lecture 21: Fundamental Theorem of Calculus

Thu 04 Nov 2021 13:03

For the duration of this lecture, $[a, b]$ will denote a compact interval in $\mathbb{R}$, principally, it is not in $\overline{\mathbb{R}}$.

Lemma 13.2. Suppose $f : [a, b] \to \overline{\mathbb{R}}$ is integrable. Then, $f = 0$ almost everywhere if and only if $\int_{[a,x]} f = 0$ for all $x \in [a, b]$.

Proof. If $f = 0$ almost everywhere, then the integral must be 0 for all $x \in [a, b]$ so the forward implication holds.
Conversely, assume $\int_{[a,x]} f = 0$ for all $x \in [a, b]$. Then, let $E = \{x \in [a, b] : f(x) > 0\}$ and assume $m(E) > 0$. Then, there is a closed set $C \subset E$ so that $m(C) > 0$. Letting $O = (a, b) \setminus C$ (an open set) we see $\int_{[a,b]} f = \int_C f + \int_O f$ and as $\int_C f > 0$ as $C \subseteq E$ with $m(C) > 0$. Hence, we find $\int_O f \neq 0$. Hence, $m(O) > 0$, and there is an interval $(c, d) \subseteq O$ so that $\int_{[c,d]} \neq 0$. Since $\int_{[a,d]=0}$ by assumption, then we find $\int_{[a,d]} f = \int_{[a,c]} f + \int_{[c,d]} f$, hence $\int_{[a,c]} f \neq 0$ ↯. $\square$

Proposition 13.3. Syppose $g : [a, b] \to \mathbb{R}$ is continuous. For every $x \in [a, b)$ and $\varepsilon > 0$ there is a δ with $0 < \delta < b - x$ such that

$$\left|\frac{1}{h}\int_{x,x+h}(g - g(x))\right| < \varepsilon \text{ for } 0 < h < \delta \ .$$

Proof. Write $g(x) = g(x)\chi_{[x,x+h]}$. Then the claim immediately follows. $\square$

Theorem 13.4 (Fundamental Theorem of Calculus I)**.** Suppose $f : [a,b] \to \overline{\mathbb{R}}$ is integrable. Then the function

$$\begin{aligned} F : [a,b] &\longrightarrow \mathbb{R} \\ x &\longmapsto F(x) = \int_{[a,x]} f \end{aligned}$$

is absolutely continuous and differentiable almost everywhere with $F' = f$ almost everywhere.

Proof. It is clear that F is absolutely continuous and differentiable almost everywhere by a result from last lecture and the fact that absolute continuity $\Rightarrow$ bounded variation $\Rightarrow$ differentiable a.e.
Moreover, we can assume $f \geq 0$, otherwise replacing f by f^+ or f^-. We can temporarily assume f is bounded (though we will later remove this requirement). Let $f(x) \leq M$ for all $x \in [a,b]$. Then, extend f, F to functions on $[a,\infty)$ by letting $f(x) = f(b)$ for all $x \geq b$. Define the following sequence of continuous functions (g_n)

$$\begin{aligned} g_n : [a,b] &\longrightarrow \overline{\mathbb{R}} \\ x &\longmapsto g_n(x) = n\left(F\left(x + \frac{1}{n}\right) - F(x)\right) \quad = n\left(\int_{a,x+\frac{1}{n}} f - \int_{a,x} f\right) \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad = n\int_{\left[x,x+\frac{1}{n}\right]} f. \end{aligned}$$

Then, we find the sequence is pointwise convergent with limit $F'(x)$ for almost every $x \in [a,b]$. Furthermore, F' is measurable and $0 \leq g_m \leq M$ for all $x \in [a,b]$. So,applying dominated convergence and the previous proposition yields g_m is dominated by M with pointwise limit F', so $F' \leq M$ almost everywhere. So, F' is integrable and for all $x \in [a,b]$ we find

$$\begin{aligned} \int_{[a,x]} F' &= \lim_{n\to\infty} \int_{[a,x]} g_n \\ &= \lim_{n\to\infty} n\left(\int_{\left[a+\frac{1}{n},x+\frac{1}{n}\right]} F - \int_{[a,x]} F\right) \\ &= \lim_{n\to\infty} n\left(\int_{\left[x,x+\frac{1}{n}\right]} F - \int_{\left[a,a+\frac{1}{n}\right]} F\right) \\ &= F(x) - F(a) \\ &= F(x). \end{aligned}$$

Now, if f was unbounded, then define the sequences (f_n) and (F_n) with

$$\begin{aligned} f_n : [a,b] &\longrightarrow \overline{\mathbb{R}} \\ x &\longmapsto f_n(x) = \inf\{f(x), n\} \\ F_n : [a,b] &\longrightarrow \overline{\mathbb{R}} \\ x &\longmapsto F_n(x) = \int_{[a,x]} f_n. \end{aligned}$$

Since $f - f_n \geq 0$, we see $F - F_n$ is increasing for each n. Hence, $F - F_n$ is differentiable almost everywhere with $(F - F_n)' \geq 0$ almost everywhere. Consequently for $x \in [a,b]$ we see

$$\int_{[a,x]} F' \geq \int_{[a,x]} F_n'$$

for all $x \in [a,b]$. Since F_n is bounded for all n , we see $\int_{[a,x]} F_n' = F_n(x)$ by the bounded case. Thus, $\int_{[a,x]} F' \geq F_n(x)$ for all $x \in [a,b]$.
Now, applying MCT, we see (f_n) is a pointwise convergent sequence of functions which are increasing the F_ns also converge pointwise to F on $[a,b]$. Hence, $\int_{[a,x]} F' \geq F(x)$ for ever $x \in [a,b]$ by passing the earlier inequality to the limit. Since f is nonnegative, we see F is increasing, so we also have $\int_{[a,x]} F' \leq F(x) - F(a) = F(x)$. Hence $\int_{[a,x]} F' = F(x)$ since

$$\int_{[a,x]} (F' - f) = \int_{[a,x]} F' - \int_{[a,x]} f = \int_{[a,x]} F' - F(x) = 0 \text{ for a.e. } x \in [a,b] .$$

□

In order to prove the other part of the fundamental theorem of calculus, we will need the following lemma:

Lemma 13.3. If the function $f : [a,b] \to \mathbb{R}$ is absolutely continuous with $f' = 0$ almost everywhere then f is a constant function.

Proof. We will show $f(c) = f(a)$ for all $c \in (a,b]$. Fix $c \in (a,b]$ and let $E = \{x \in (a,c) : f' \text{ exists at } x, f'(x) = 0\}$.
By assumption, $m(E) = c - a > 0$, hence for $\varepsilon > 0$ choose $\delta > 0$ such that absolute continuity holds. For each $x \in E$ and $k > 0$, we see there is an $h \in (0,k)$ with either $[x, x+h] \subseteq [a,c]$ and $|f(x+h) - f(x)| < \varepsilon h$ or $[x-h, x] \subseteq [a,c]$ and $|f(x-h) - f(x)| < \varepsilon h$ (or both). Then, the collection $\mathscr{C}$ of these intervals for all $k > 0$ and $x \in E$ is a vitali covering of E. By the Vitali covering lemma, we find a finite disjoint collection $\{[x_k, y_k] \in \mathscr{C} : 1 \leq k \leq n\}$ so that $V = \bigcup_{k=1}^{N} [x_k, y_k]$ has $m(E \setminus V) < \delta$. Reindex these intervals such that $x_k < x_{k+1}$ for all k and let $y_0 = a$, $x_{n+1} = c$. Then, we see

$$a = y_0 \leq x_1 < y_1 < x_2 < y_2 < \ldots < x_n < y_n \leq x_{n+1} = c.$$

Hence, the set $P = \{x_i : 1 \leq i \leq n+1\} \cup \{y_i : 1 \leq i \leq n+1\}$ is a partition of $[a, c]$. Since

$$\sum_{k=1}^{n} (y_k - x_k) = m(V) > m(E) = c - a - \delta$$

we see the leftover pieces

$$\sum_{k=0}^{n} (x_{k+1} - y_k) \leq m(E \setminus V) < \delta.$$

Since f is absolutely continuous, we see $\sum_{k=0}^{n} |f(x_{k+1}) - f(y_k)| < \varepsilon$. Consequently,

$$\begin{aligned} |f(c) - f(a)| &\leq \sum_{k=1}^{n} |f(y_k) - f(x_k)| + \sum_{k=0}^{n} |f(x_{k+1} - f(y_k))| \\ &< \sum_{k=1}^{n} \varepsilon (y_k - x_k) + \varepsilon \\ &\leq \varepsilon (c - a) + \varepsilon \end{aligned}$$

for all $\varepsilon > 0$, so we see $f(c) - f(a) = 0$ for all $c \in (a, b]$ and the claim follows. $\square$

Theorem 13.5 (Fundamental Theorem of Calculus II)**.** Suppose the function $F : [a, b] \to \mathbb{R}$ is absolutely continuous. Then, F is differentiable almost everywhere and its derivative, F', is integrable with

$$\int_{[a,x]} F' = F(x) - F(a)$$

for all $x \in [a, b]$.

Proof. Since F is absolutely continuous, it is of bounded variation, so there are two increasing functions, $T, S : [a, b] \to \mathbb{R}$ with $F = T - S$. Moreover, the derivatives T', S' exist almost everywhere and are integrable.
Hence, F' exists almost everywhere and $F' = T' - S'$ almost everywhere, so it is integrable as well.
Then, letting $G(x) = \int_{[a,x]} F'$. We see G is absolutely continuous, so $F - G$ must be absolutely continuous. Then, by the FTC part 1, we see $(F - G)'$ exists almost everywhere and $(F - G)'(x) = 0$ for almost every $x \in [a, b]$. Hence $F - G$ is a constant function. So, we see $F(x) - G(x) = F(x) - \int_{[a,x]} F' = F(a)$ by letting $x = a$. $\square$