Analysis 3

Notes by Lev Leontev
Professor: Igors Gorbovickis

May 4, 2023

Contents

1 Measure

1.1 Introduction

We want to generalize the notion of the *length* towards all the subsets of $\mathbb{R}$. Such a generalized function is usually called *measure*. But, unfortunately, such a function does not exist.

Theorem 1. There exist no such function $\mu : 2^{[0,1]} \to [0, +\infty)$ that satisfies the following properties:

1. The function is non-negative;
2. It's countably additive;
3. It's monotonic: the measure of a subset is not greater than the entire set;
4. Translation does not change the measure;
5. The measure of the unit interval is 1.

Proof. First, several definitions:

Step 1. Let's define the following equivalence relation: if x, y are from the unit interval, we'll say that $x \sim y$ if $x - y \in \mathbb{Q}$.

Step 2. Let's choose $N \subset [0, 1/3]$ such that it contains *precisely one* element from each equivalence class. (Such an N exists if the axiom of choice holds true).

Step 3. For all $r \in \mathbb{Q}$ define $N_r = N + r$.

Claim 1. The sets N_R are congruent to N and are pairwise disjoint.

Proof. The sets are congruent by definition. Let's prove that they are pairwise disjoint.

Assume that $x \in N_{r_1} \cap N_{r_2}$ for some $r_1, r_2 \in \mathbb{Q}$. Then $x - r_1 \in N$, $x - r_2 \in N$, but $(x - r_1) \sim (x - r_2) \implies r_1 = r_2$.

Claim 2.

$$\left[\frac{1}{3}, \frac{2}{3}\right] \in \bigcup_{r \in \mathbb{Q} \cap [0,2/3]} N_r$$

Proof. If $x \in [1/3, 2/3]$, then $\exists!\, y \in N$ such that $x = y + q$ for some $q \in \mathbb{Q}$, as N contains exactly one representative from each of the equivalence classes. It is easy to see that such $q \in [0, 2/3]$.

We arrive at the following conclusion:

$$\frac{1}{3} = \mu\big([1/3, 2/3]\big) \leqslant \mu\Big(\bigcup_{r \in \mathbb{Q} \cap [0,2/3]} N_r\Big) = \sum_{r \in \mathbb{Q} \cap [0,2/3]} \mu(N_r) \leqslant 1$$

What is $\mu(N)$ then? If $\mu(N) = 0$, then

$$\mu\Big(\bigcup_{r \in \mathbb{Q} \cap [0,2/3]} N_r\Big) = \sum 0 = 0$$

If $\mu(N) = \varepsilon > 0$, then the sum is $+\infty$. But it's supposed to be in $[1/3, 1]$?! $\square$

Consequence. We cannot generalize the notion of length to all subsets of real numbers.

1.2 Lebesgue Outer Measure

***Definition* 1.** If $I \subset \mathbb{R}$ is an interval, then $l(I) =$ the length of I. If I is unbounded, then $l(I) = \infty$.

***Definition* 2** (Outer Measure)**.**

$$m^* : 2^{\mathbb{R}} \to [0, +\infty]$$
$$m^*(A) = \inf\Big\{ \sum_{j=1}^{\infty} l(I_j) \mid I_j \text{ — open intervals, } A \subseteq \bigcup_{j=1}^{\infty} I_j \Big\}$$

In words, it's the infimum of all *countable* covers of A. (A countable sum either converges or diverges to infinity).

Remark. This is certainly not a measure — otherwise, it would contradict Theorem 1.

Example. If A is countable, then $m^*(A) = 0$.

Proof. Let's choose an arbitrary $\varepsilon > 0$ and prove that $m^*(A) \leqslant 2\varepsilon$. Let's choose a cover of the points with segments of lengths ε, $\varepsilon/2$, $\varepsilon/2^2$, and so on. Then

$$m^*(A) = \inf\{\dots\} \leqslant \varepsilon + \frac{\varepsilon}{2} + \frac{\varepsilon}{2^2} + \dots = 2\varepsilon$$

□

Proposition 1. If A is an interval, then $m^*(A) = l(A)$.

Proof. a) A is a closed interval, $A = [a, b]$.

1. $m^*(A) \leqslant b - a$. To prove this, we can cover A with a single interval:

$$(a - \varepsilon, b + \varepsilon) \implies \sum l(I_j) = b - a + 2\varepsilon$$

Now take $\varepsilon \to 0$.

2. $m^*(A) \geqslant b - a$. Suppose we an infinite cover of A by open intervals. Since A is a compact set, we can choose a finite subcover. The case of a finite cover with open intervals is simple. We can prove it as follows: if we have two intersecting open intervals, we can replace them with a single interval of a lesser length. Then we can continue this process using induction.

b) If A is unbounded, then all of the covers would have infinite sum, and thus the infimum will be infinite as well.

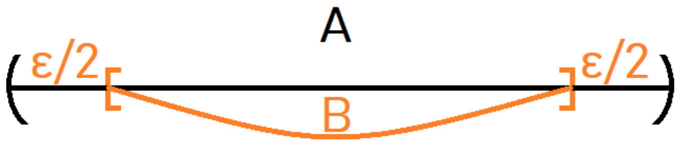
c) If A is an open or semiclosed interval, we can approximate it from both sides by closed intervals. Let's denote the closure of A by $\bar{A}$. Since we're adding points, the Outer Measure will not decrease:

$$A \subset \overline{A} \implies m^*(A) \leqslant m^*(\overline{A}) = l(a)$$

Now suppose we have a closed interval B strictly inside A. Then we get

$$m^*(A) \geqslant m^*(B) = l(B) = l(A) - \varepsilon$$

Now take $\varepsilon \to 0 \implies m^*(A) \geqslant l(A)$.



$\square$

Lemma. m^* is translation-invariant.

Proof. If we translate the set, we can translate all of its covers as well. Since translating an interval does not change its length, the lengths of the covers won't change either. $\square$

Proposition 2 (Countable subadditivity)**.** For any countable collection of sets $\{E_k\}_{k=1}^{\infty}$ we have

$$m^*\left(\bigcup_{k=1}^{\infty} E_k\right) \leqslant \sum_{k=1}^{\infty} m^*(E_k)$$

Remark. We don't ask for the sets E_k to be disjoint. If we proved that we have an equality sign for the disjoint case, we would have proved that m^* is a measure, which we proved does not exist in Theorem 1.

Proof. Choose open intervals $I_{k,i}$, such that

$$E_k \subset \bigcup_{i=1}^{\infty} E_{k,i} \ (E_{k,i} \text{ are a cover of } E_k)$$

and

$$\sum_{i=1}^{\infty} l(I_{k,i}) < m^*(E_k) + \frac{\varepsilon}{2^k}$$

Such intervals exist from the definition of the infimum.

On the other hand, $\{I_{k,i} \mid 1 \leqslant k, i < \infty\}$ covers each of the E_k, and thus it's a cover of $\cup_{k=1}^{\infty} E_k$. Then

$$m^*\left(\bigcup_{k=1}^{\infty} E_k\right) \overset{\text{it's a cover}}{\leqslant} \sum_{1 \leqslant k,i < \infty} l(I_{k,i}) < \sum_{k=1}^{\infty} m^*(E_k) + \varepsilon\left(\frac{1}{2} + \frac{1}{4} + \dots\right) = \sum_{k=1}^{\infty} m^*(E_k) + \varepsilon$$

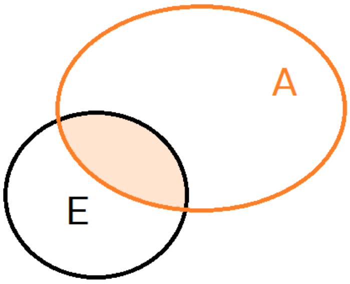
Now take $\varepsilon \to 0$. $\square$

Remark. Here we assume that all of the E_k have finite outer measures. Otherwise, both of the sides of the inequality would diverge to infinity, and we get $\infty \leqslant \infty$ which is "true".

1.3 The σ-algebra of Lebesgue-measurable sets.

***Definition* 1.** A set E is (Lebesgue) measurable if for any set A,

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^C) \qquad E^C = \mathbb{R} \setminus E$$



The set E "splits" A into two parts

***Remark*.** We already have the $\leqslant$ sign from countable subadditivity.

***Remark*.** Motivation: If $A \cap B = \varnothing$ and A (or B) is measurable, then

$$m^*(A \cup B) = m^*\big((A \cup B) \cap A\big) + m^*\big((A \cup B) \cap A^C\big) = m^*(A) + m^*(B)$$

Proposition 1. If $m^*(E) = 0$, then E is measurable.

Proof. For all A we have:

$$m^*(A \cap E) \leqslant m^*(E) = 0 \implies m^*(A \cap E) = 0$$
$$m^*(A) \geqslant m^*(A \cap E^C) = m^*(A \cap E) + m^*(A \cap E^C)$$

As we noted earlier, the inequality in the other side follows from countable subadditivity. □

Proposition 2. If $E_1, \dots, E_n$ are measurable, then $\cup_1^n E_k$ is measurable.

Proof. Case $n = 2$: for all A we have

$$m^*(A) = m^*(A \cap E_1) + m^*(A \cap E_1^C) =$$
$$= m^*(A \cap E_1) + m^*\big((A \cap E_1^C) \cap E_2\big) + m^*\big((A \cap E_1^C) \cap E_2^C\big) = (*)$$
$$X := A \cap E_1,\ Y := (A \cap E_1^C) \cap E_2,\ Z := (A \cap E_1^C) \cap E_2^C$$

With Venn diagrams it's possible to prove that $Z = A \cap (E_1 \cup E_2)^C$, $X \cup Y = A \cap (E_1 \cup E_2)$. Now let's apply countable subadditivity to X and Y. Then we get:

$$(*) \geqslant m^*\big(A \cap (E_1 \cup E_2)\big) + m^*\big(A \cap (E_1 \cup E_2)^C\big)$$

Yet again, the inequality in the other side follows from countable subadditivity.

Induction step: Apply case $n = 2$ to the sets $\cup_1^{n-1} E_k$, E_n. □

Definition 2 (Algebra)**.** Let X be a non-empty set. $\Omega \subset 2^X$ is an algebra, if:

1. $X \in \Omega$;
2. Ω is closed under the formation of complements in X and *finite* unions.

Remark. It follows that Ω is also closed under intersections:

$$\left(X_1^C \cup \cdots \cup X_n^C\right)^C = X_1 \cap \cdots \cap X_n$$

Definition 3 (σ-algebra)**.** Let X be a non-empty set. $\Omega \subset 2^X$ is a σ-algebra, if:

1. $X \in \Omega$;
2. Ω is closed under the formation of complements in X and *countable* unions.

Remark. Every σ-algebra is an algebra, but not vice versa.

Corollary 1. The collection $\mathcal{M}$ of all measurable subsets of $\mathbb{R}$ is an algebra.

Proof. For the proof, we'll need to show that:

1. $\mathbb{R}$ is measurable.
$$m^*(A) = m^*(A \cap \mathbb{R}) + m^*(A \cap \mathbb{R}^C) = m^*(A) + m^*(\varnothing)$$
2. It is closed under complements. It follows from the symmetry of the definition of a measurable set.
3. It is closed under unions. We have already proved this one.

□

Proposition 3. $\{E_k\}_1^n$ — disjoint measurable sets. Then for every set A

$$m^*\left(A \cap \left[\bigcup_1^n E_k\right]\right) = \sum_1^n m^*(A \cap E_k)$$

In particular, for $A = \mathbb{R}$ we have

$$m^*\left(\bigcup_1^n E_k\right) = \sum_1^n m^*(E_k)$$

Proof. Induction on n.

Base $n = 1$ is obvious.

Step $n-1 \to n$. Take $\hat{A} := A \cap \left[\bigcup_1^n E_k\right]$. Then

$$\hat{A} \cap E_n = A \cap E_n$$

We also have

$$\hat{A} \cap E_n^C = A \cap \left[\bigcup_1^{n-1} E_k\right]$$

That is true, as intersecting with E_n^C is equivalent to subtracting E_n from $\hat{A}$, and since $\{E_k\}$ are disjoint, no other parts of $\hat{A}$ except E_n will be removed. Then:

$$m^*(\hat{A}) \overset{E_n \text{ is measurable}}{=\!=\!=} m^*(\hat{A} \cap E_n) + m^*(\hat{A} \cap E_n^C) =$$
$$= m^*(A \cap E_n) + m^*\left(A \cap \left[\bigcup_1^{n-1} E_k\right]\right) \overset{\text{induction}}{=} m^*(A \cap E_n) + \sum_1^{n-1} m^*(A \cap E_k)$$

□

Proposition 4. The union of a countable collection of measurable sets is the union of a countable collection of *disjoint* measurable sets.

Proof. If $A = \cup_1^\infty A_k$, define $\hat{A}_1 := A_1$ and $\hat{A}_k := A_k \setminus \cup_1^{k-1} A_j$. As $\mathcal{M}$ is an algebra, all $\hat{A}_k$ are measurable, and $A = \sqcup_1^\infty \hat{A}_k$, which is what we wanted. □

Theorem 1. $\mathcal{M}$ is a σ-algebra.

Proof. We need to show that if all $\{E_k\}_1^\infty$ are measurable sets, then $E = \cup_1^\infty E_k$ is measurable. By Proposition 4, without the loss of generality, assume that E_k are all pairwise disjoint. Let $F_n := \cup_1^n E_k$, then $F_n \in \mathcal{M}$ (as a finite union). As $F_n \subset E$, we have $E^C \subset F_n^C$.

Let A be any set. Then:

$$m^*(A) = m^*(A \cap F_n) + m^*(A \cap F_n^C) \geqslant m^*(A \cap F_n) + m^*(A \cap E^C) \overset{\text{Proposition 3}}{=}$$
$$= \sum_1^n m^*(A \cap E_k) + m^*(A \cap E_C)$$

Now take $n \to \infty$:

$$m(A) \geqslant \sum_1^\infty m^*(A \cap E_k) + m^*(A \cap E_C) \overset{\text{countable subadditivity}}{\geqslant} m^*(A \cap E) + m^*(A \cap E^C)$$

Now we have the inequality in the difficult direction. The inequality in the other direction is obvious (again, from countable subadditivity). □

Proposition 5 (Countable additivity)**.** If $\{E_k\}_1^\infty \subset \mathcal{M}$ — collection of disjoint sets, then $\cup_1^\infty E_k \in \mathcal{M}$ and

$$m^*\Big(\bigcup_1^\infty E_k\Big) = \sum_1^\infty m^*(E_k)$$

Proof. We know that:

1. $$m^*\Big(\bigcup_1^\infty E_k\Big) \leqslant \sum_1^\infty m^*(E_k) \text{ (countable } \textit{sub}\text{additivity)}$$

2. $$m^*\Big(\bigcup_1^\infty E_k\Big) \geqslant m^*\Big(\bigcup_1^n E_k\Big) \overset{\text{Proposition 3}}{=} \sum_1^n m^*(E_k)$$

   Take $n \to \infty$, then

   $$m^*\Big(\bigcup_1^\infty E_k\Big) \geqslant \sum_1^\infty m^*(E_k)$$

   Which is what we wanted.

□

***Definition* 4.** The restriction of m^* on $\mathcal{M}$ is called the Lebesgue measure and denoted by m.

$$m(E) := m^*(E) \quad \forall E \in \mathcal{M}$$

***Definition* 5.** If X is a non-empty set and $\mathcal{A}$ is a σ-algebra on X, then any function $\mu : \mathcal{A} \to [0, +\infty]$ is called the measure on $(X, \mathcal{A})$, if:

1. $\mu(\varnothing) = 0$.
2. μ is countable additive.

***Definition* 6** (Measurable space)**.** A *measurable space* is a tuple $(X, \mathcal{A})$, where:

1. X is a set.
2. $\mathcal{A}$ is a σ-algebra on X.

***Definition* 7** (Measure space)**.** A *measure space* is a triple $(X, \mathcal{A}, \mu)$, where:

1. X is a set.
2. $\mathcal{A}$ is a σ-algebra on X.
3. μ is a measure on $(X, \mathcal{A})$.

Example 1. $\{\varnothing, X\}$ is a σ-algebra. Any μ, such that $\mu(\varnothing) = 0$ and $\mu(X) \geqslant 0$ will be a measure.

Example 2. 2^X is a σ-algebra. We can have the following measures:

a) $\mu(E) = |E|$ is called a *counting measure*. Here $|E|$ denotes the cardinality of E (number of elements in E).

b) δ-measure (also called Dirac measure):

$$\mu(E) = \begin{cases} 1, & 0 \in E \\ 0, & \text{otherwise} \end{cases}$$

1.4 Continuity of measure

***Definition* 1.** A countable collection of sets $\{E_k\}_{k=1}^{\infty}$ is called *ascending* if $E_k \subset E_{k+1}$.

***Definition* 2.** A countable collection of sets $\{E_k\}_{k=1}^{\infty}$ is called *descending* if $E_k \supset E_{k+1}$.

Theorem 1 (Continuity of measure)**.**

1. If $\{A_k\}_{k=1}^{\infty} \subset \mathcal{A}$ and the sequence is ascending, then

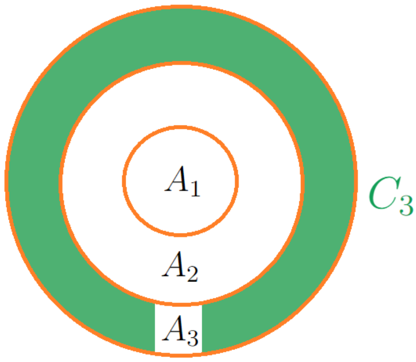
$$\mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \lim_{k\to\infty} \mu(A_k)$$

2. If $\{B_k\}_{k=1}^{\infty} \subset \mathcal{A}$, the sequence is descending and $\mu(B_1) < \infty$, then

$$\mu\Big(\bigcap_{k=1}^{\infty} B_k\Big) = \lim_{k\to\infty} \mu(B_k)$$

Proof. 1. Let $C_k \coloneqq A_k \setminus A_{k-1}$. Then we have:

$$\mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \mu\Big(\bigsqcup_{k=1}^{\infty} C_k\Big) = \sum_{k=1}^{\infty} \mu(C_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(C_k) = \lim_{n\to\infty} \mu(A_n)$$



2. Let $D_k \coloneqq B_1 \setminus B_k$. Since B_k is descending, it follows that D_k is an ascending sequence. Then from the part 1 of the theorem it follows that:

$$\mu\Big(\bigcup_{k=1}^{\infty} D_k\Big) = \lim_{k\to\infty} \mu(D_k) \qquad \bigcup_{k=1}^{\infty} D_k = B_1 \setminus \bigcap_{k=1}^{\infty} B_k$$

$$\mu\Big(B_1 \setminus \bigcap_{k=1}^{\infty} B_k\Big) = \lim_{k\to\infty} \big(\mu(B_1) - \mu(B_k)\big) = \mu(B_1) - \lim_{k\to\infty} \mu(B_k)$$

$$\mu\Big(B_1 \setminus \bigcap_{k=1}^{\infty} B_k\Big) = \mu(B_1) - \mu\Big(\bigcap_{k=1}^{\infty} B_k\Big) \implies \mu\Big(\bigcap_{k=1}^{\infty} B_k\Big) = \lim_{k\to\infty} \mu(B_k)$$

□

Definition 3. We say that a statement (property) holds for *almost all $x \in X$ with respect to a measure μ*, if $\exists N \in \mathcal{A}$, such that $\mu(N) = 0$ and the statement (property) holds for all $x \in X \setminus N$.

Lemma (Borel-Cantelli)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $\{E_k\}_{k=1}^{\infty} \subset \mathcal{A}$ and $\sum_{k=1}^{\infty} \mu(E_k) < \infty$. Then *almost all* $x \in X$ belong to at most finitely many E_k.

Proof. Let $B_n = \cup_{k=n}^{\infty} E_k$. It's easy to see that B_k is a descending measure. At the same time,

$$\mu(B_1) = \mu\Big(\bigcup_{k=1}^{\infty} E_k\Big) \leqslant \sum_{k=1}^{\infty} \mu(E_k) < \infty$$

By definition of B_n, $\cap_{n=1}^{\infty} B_n$ contains all the points that are contained in infinitely many E_k's. But, by continuity of measure for $\{B_n\}_{n=1}^{\infty}$ we have:

$$\mu\Big(\bigcap_{n=1}^{\infty} B_n\Big) = \lim_{n\to\infty} \mu(B_n) = \lim_{n\to\infty} \Big(\bigcup_{k=n}^{\infty} E_k\Big) \leqslant \lim_{n\to\infty} \sum_{k=n}^{\infty} \mu(E_k) = 0$$

□

1.5 How large is the Lebesgue σ-algebra $\mathcal{M}$?

Proposition 1. Every interval is Lebesgue-measurable.

Proof. Proof idea:

$$E \in \mathcal{M} \iff \forall A : m(A) = m(A \cap E) + m(A \cap E^C)$$

Assume $E = (-\infty, a)$. If we prove that such intervals lie in $\mathcal{M}$, then we'll prove everything (since $\mathcal{M}$ is a σ-algebra). We already have $m(A) \leqslant m(A \cap E) + m(A \cap E^C)$ from countable subadditivity.

Let's assume $a \notin A$ (since removing one point does not change the measure). Every cover of A can be split into two covers with the same sum of interval lengths: of $A \cap (-\infty, a)$ and $A \cap (a, +\infty)$. Every interval in those covers, that contains a, can be split into two. Therefore, from the definition of Lebesgue measure, $m(A) \geqslant m(A \cap E) + m(A \cap E^C)$, so we've proved the inequality in both sides. □

***Definition* 1.** For any $\mathcal{X} \in 2^{\mathbb{R}}$ let $\mathcal{A}(\mathcal{X})$ be the smallest σ-algebra containing $\mathcal{X}$.

Lemma. $\mathcal{A}(\mathcal{X})$ always exists and is the intersection of all σ-algebras containing $\mathcal{X}$.

Proof. We have to prove that if we intersect a bunch σ-algebras, we still get a σ-algebra.

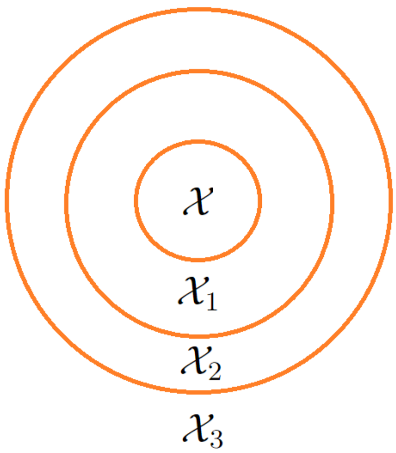
1. Such an intersection is closed under complements: if a set belongs to the intersection of σ-algebras, then it belongs to each of the σ-algebras, then its complement belongs to each of the σ-algebras, and thus its complement belongs to the intersection of σ-algebras.

2. In a similar way, such an intersection is closed under countable unions: if a number of sets all belong to the intersection of σ-algebras, then they all belong to each of the σ-algebras, then their countable union belongs to each of the σ-algebras, and their countable union belongs to the intersection of σ-algebras.

□

Remark. We can try to construct $\mathcal{A}(\mathcal{X})$ in a different way. Say, $\mathcal{X}$ is not a σ-algebra. Let's enlarge it: first by including all the complements. Then let's enlarge it by all countable unions. Let's call such a set $\mathcal{X}_1$. But after such operation, $\mathcal{X}_1$ may be non-closed under complements. So we repeat such a procedure.

And, in general: $\mathcal{X}_{n+1}$ is obtained from $\mathcal{X}_n$ is obtained by including into $\mathcal{X}_n$ all complements of the sets from $\mathcal{X}_n$ and then including all countable unions of the obtained sets.



It is tempting to think that $\cup_1^\infty \mathcal{X}_i$ is $\mathcal{A}(\mathcal{X})$. Is it true? No, not necessarily. If the sequence $\{\mathcal{X}_i\}$ eventually stabilizes, then such a construction works. Let's now assume that every next $\mathcal{X}_i$ is larger than the previous one. Then we can take A from $\mathcal{X}$, A_1 from $\mathcal{X}_1 \setminus \mathcal{X}$, A_2 from $\mathcal{X}_2 \setminus \mathcal{X}_1$, and so on.

Now let's look at $\cup_1^\infty A_i$. As a countable union, it must be contained in $\mathcal{A}(\mathcal{X}) = \cup_1^\infty \mathcal{X}_i$, thus, there exist an n, such that $\cup_1^\infty A_i \in \mathcal{X}_n$. But $A_{n+1} \in \mathcal{X}_{n+1} \setminus \mathcal{X}_n$?!

***Definition* 2** (Topological space)**.** A *topological space* is a set X and a collection of subsets O of X (called *open sets*), such that $\varnothing, X \in O$, and:

1. A union of (possibly infinitely many) sets from O is in O.

2. The intersection of finitely many sets from O is in O.

The complements of open sets are called *closed sets*.

***Definition* 3.** A function $f : X \to Y$ between two topological spaces is *continuous* if the preimage of every open set is open.

Remark. It is possible to check that for $\mathbb{R}$ this definition is equivalent to the usual one.

***Definition* 4** (Borel σ-algebra)**.** For a topological space X its *Borel σ-algebra* $\mathcal{B}_X$ is the smallest σ-algebra on X that contains all open sets.

Remark. If it's obvious from the context which set we are talking about, we will just write $\mathcal{B}$ (without a subscript).

Theorem 1. $\mathcal{B}_\mathbb{R} \subset \mathcal{M}$ (all of the sets in $\mathcal{B}_\mathbb{R}$ are measurable).

Proposition 2. $\mathcal{B}$ is the smallest σ-algebra that contains all open intervals.

If we prove the proposition, the theorem will follow easily. We know that all the intervals are Lebesgue-measurable. We know that the Lebesgue-measurable sets ($\mathcal{M}$) are a σ-algebra. Thus, if we take the smallest σ-algebra that contains all open intervals, it will be a subset of $\mathcal{M}$.

Proof of Proposition 2. We will prove that every open set $O \subset \mathbb{R}$ is a finite or countable union of open intervals.

For every point $x \in O$ let I_x be the largest open interval, such that $x \in I_x$ and $I_x \subset O$. It exists as a union of all such intervals. Since O is open, x lies in O with an open neighborhood, thus, I_x is non-empty.

$$\forall x \in O : x \in I_x \implies O = \bigcup_{x \in O} I_x$$

Let's prove that $I_x \cap I_y \neq \varnothing \implies I_x = I_y$. If the intervals around x and y intersect, then $I_x \cup I_y$ is an interval as well, and $I_x \cup I_y \in O$ as $I_x \in O$ and $I_y \in O$. Since I_x and I_y are the largest such intervals, it follows that $I_x = I_x \cup I_y = I_y$.

Let's say that two points x and y are equivalent if $I_x = I_y$. Since there's a lot of same intervals in $O = \cup_{x \in O} I_x$, we can take just a single point from every equivalence class and still get O as a union. Particularly, every open interval contains at least one rational point (as rational numbers are dense). Therefore, there's a rational point in every equivalence class. Thus,

$$O = \bigcup_{x \in O \cap \mathbb{Q}} I_x$$

Since the set of rational numbers is countable, we have represented O as a countable union of open intervals, which is what we wanted. □

***Definition* 5.** A topological space is called *separable*, if it contains a countable dense subset.

Remark. We have proved that the Lebesgue measure exists on $\mathcal{B}_\mathbb{R}$, so we have a lot of measurable sets.

Remark. The Lebesgue measure can be generalized to $\mathbb{R}^n$.

1.6 Other critirea of measurability for Lebesgue measure

As we remember, the definition of a measurable set is difficult to check. Thus, we would like to have better critirea.

Theorem 1. $E \subset \mathbb{R}$ is Lebesgue measurable if and only if one of the following holds:

1. For every $\varepsilon > 0$ there exists an open set O, such that $E \subset O$ and $m^*(O \setminus E) < \varepsilon$.
2. There exists a G_δ-set G, such that $E \subset G$ and $m^*(G \setminus E) = 0$.
   (A G_δ-set is a countable intersection of open sets.)
3. For every $\varepsilon > 0$ there exists a closed set F, such that $F \subset E$ and $m^*(E \setminus F) < \varepsilon$.
4. There exists a F_σ-set F, such that $F \subset E$ and $m^*(E \setminus F) = 0$.
   (A F_δ-set is a countable union of closed sets.)

Proof.

- E is measurable $\implies$ 1.

  If $m^*(E) < \infty$, then from the definition of m^* we can find O — a finite union of open intervals, such that $m^*(O) < m^*(E) + \varepsilon$. Since O is an open set, it's measurable (as we proved earlier). Therefore, both E and O are measurable, thus

$$\varepsilon > m^*(O) - m^*(E) \overset{E,O\in\mathcal{M}}{=} m^*(O \setminus E)$$

  If $m^*(E) = \infty$, let's split the set E into a countable number of sets with finite measure. For example, by splitting the real line into segments of length 1. So, $E = \cup_1^\infty E_k$, where $m^*(E_k) < \infty$. Then let's use geometrically decreasing ε's for the covers of each E_k: $\varepsilon/2$ for E_1, $\varepsilon/4$ for E_2, and so on. When we sum up the inequalities, the fractions will sum up to ε. So, we obtained our O, now continue like in the previous case.

  ***Definition* 1.** A measure μ on X is called σ-*finite*, if $X = \cup_1^\infty X_k$ and $\mu(X_k) < \infty$ for all k.

  In words: if there exists a subdivision of X into a countable number of set of finite measure.

- $1 \implies 2$.

  From 1, $\forall k \in \mathbb{N},\ \exists O_k$ — open, such that $E \subset O_k$ and $m^*(O_k \setminus E) < \frac{1}{k}$. Now let's take

$$G := \bigcap_1^\infty O_k \implies \forall k:\ m^*(G \setminus E) \leqslant m^*(O_k \setminus E) < \frac{1}{k} \implies m^*(G \setminus E) = 0$$

- $2 \implies E$ is measurable.

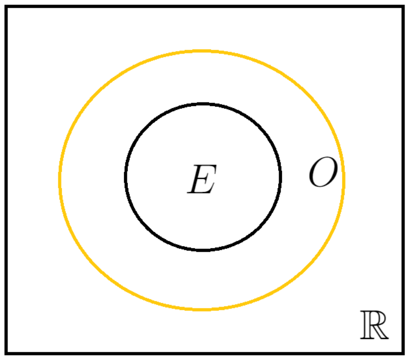
  G is a G_δ-set. As a countable intersection of open sets, it's in Borel σ-algebra, and thus is Lebesgue-measurable. $m^*(G \setminus E) = 0$, then $G \setminus E$ is Lebesgue-measurable, then $E = G \setminus (G \setminus E)$ is measurable as a difference of two measurable sets.

- $3 \iff 1,\ 4 \iff 2$.

  If we assume that 3 holds for E, then, if we take $O = \overline{F}$, 1 will hold for $\overline{E}$. Therefore, $\overline{E}$ is measurable, then E is measurable (as $\mathcal{M}$ is a σ-algebra).

  If 1 holds for E, then E is measurable, then $\overline{E}$ is measurable, then 1 holds for $\overline{E}$. Now take $F = \overline{O}$, therefore, 3 holds for E.

  In the same way, 2 and 4 are equivalent as well.



□

Theorem 2. For every $E \in \mathcal{M}$ with $m(E) < \infty$ and for every $\varepsilon < \infty$ there exists an infinite disjoint collection of open intervals $\{I_k\}_1^n$, such that $O = \cup_{k=1}^n I_k$ and $m(E \triangle O) < \varepsilon$.

(Here $\triangle$ is the symmetric difference of two sets).

Proof. From part 1 of the previous theorem, we can take such an open set U, that $E \subset U$ and $m(U \setminus E) < \varepsilon/2$.

As we proved earlier, we can represent U as a countable union of disjoint open intervals I_k. Then:

$$\forall n : \bigcup_1^n I_k \subset U \implies \forall n : \sum_1^n m(I_k) \leqslant m(U) < \infty$$

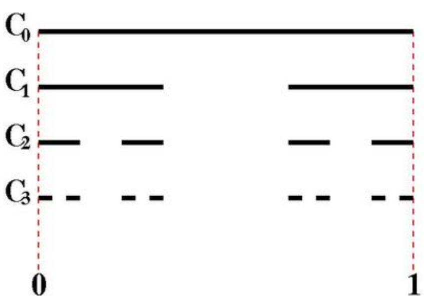
Now take n, such that $\sum_{n+1}^{\infty} m(I_k) < \varepsilon/2$, and put $O := \cup_{k=1}^n I_k$. Then $m(O \setminus E) < \varepsilon/2$ and $m(E \setminus O) < \varepsilon/2$, therefore, the measure of the symmetric difference is less that ε. □

1.7 Cantor set

Questions:

1. If $m(A) = 0$, is A countable?
2. We know that $\mathcal{B} \subset \mathcal{M}$. Is this inclusion proper?

***Definition* 1** (Cantor set)**.** Let's take $[0, 1]$, split it into three parts and remove the middle part. Then continue such process. The *Cantor set* is the set $C := \cap_0^{\infty} C_k$.



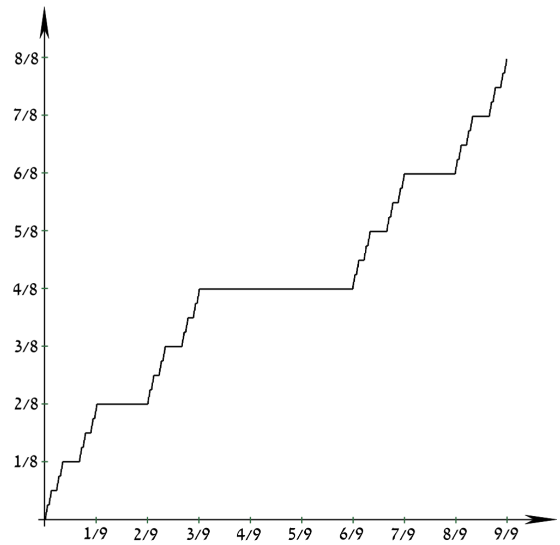
Cantor set illustration from here.

*Remark***.** The measure of C_n is $(2/3)^n = 0$. From the continuity of measure, the measure of the intersection is the limit of measures of individual sets, and thus $m(C) = 0$.

Remark. The Cantor set is countable, because if we have a sequence of zeros and ones, we can traverse down-left on 0 and down-right on 1. The intersection of the corresponding intervals will be a single point of C. So, there's a bijection between C and $\{0,1\}^{\mathbb{N}}$, therefore, C is indeed uncountable.

Remark. Usually, when we remove the middle intervals on each step, we keep the end points. If we choose to remove the end points, we essentially remove the points that correspond to sequences that end with an infinite sequence of 0's or an infinite sequence of 1's. It's clear that not all sequences of 0's and 1's are like that, so there is going to be plenty of points left in the Cantor set, anyway.

Definition 2 (The Cantor function)**.** The Cantor function $\varphi : [0,1] \to [0,1]$ is defined as follows. Let's first take the unit interval $[0,1]$, split it in three, and define φ to be $\frac{1}{2}$ on $[1/3, 2/3]$. Then let's continue the same with $[0,1/3]$ and $[2/3,1]$, and so on.



Now we have defined φ for all points on the Cantor set. Here's how we'll define it on all others:

$$O \coloneqq [0,1] \setminus C$$
$$\forall x \in [0,1] \setminus C : \ \varphi(x) \coloneqq \sup\{\varphi(t) \mid t \in C \cap [0,x)\}$$

Proposition 1. φ is increasing, continuous, surjective ($[0,1]$ acts on $[0,1]$), φ' exists for open set O of measure 1, $\varphi'\Big|_O \equiv 0$.

Proof. φ is increasing (that's clear), therefore, if φ is discontinuous, then it has a jump discontinuity and there will be and interval, say, $I \subset [0,1]$, such that $\varphi([0,1]) \cap I = \varnothing$. But

$$\varphi([0,1]) \supset \left\{ \frac{m}{2^k} \;\middle|\; k \in \mathbb{N}, m \in [0, 2^k] \right\} = K$$

And K is dense in $[0,1]$, therefore, $K \cap I \neq \varnothing$, which is a contradiction. □

Remark. The Cantor function is a source of a lot of counterexamples. We are going to use it to answer the second question from earlier.

Definition 3.

$$\psi : [0,1] \to [0,2], \ \psi(x) = \varphi(x) + x$$

ψ is strictly increasing, continuous, surjective.

Proposition 2. 1. ψ maps C onto a measurable set of measure 1.

2. ψ maps some subset of C onto a non-measurable set.

Proof. 1. $m(C) = 0$, thus, $m(O) = m([0,1] \setminus C) = 1$. We know that φ is constant on every part of O. Therefore, ψ looks like x on every part of O, and there's a countable number of such intervals. Therefore, O is mapped to a set of measure 1. Thus, $m(\psi(C)) = m(\psi([0,1])) - m(\psi(O)) = 2-1 = 1$.

2. Since $\psi(C)$ has measure $1 > 0$, by the second homework, there exists a non-measurable set $N \subset \psi(C)$. Now take $\psi^{-1}(N)$.

$\square$

Corollary 1. There is a (measurable) subset of C that is not Borel.

Remark. Since $m(C) = 0$, the outer measure of each subset of C is also 0, thus, every subset of C is measurable. So, the "measurable" part of the corollary is obvious.

Proposition 3. If $f : E \to \mathbb{R}$ is continuous, $E \in \mathcal{M}$, then the preimage of every Borel set is measurable.

$$\forall B \in \mathcal{B} : f^{-1}(B) = \{x \in E \mid f(x) \in B\} \in \mathcal{M}$$

Let's show how we can prove the corollary if we prove the proposition:

Proof of Corollary 1. $\psi^{-1} : [0,2] \to [0,1]$ is continuous and bijective. Let's define $\tilde{C} = \psi^{-1}(N)$. Since $N \subset C$, we have $\tilde{C} \subset C$. Assume that $\tilde{C}$ is a Borel set, and thus it's measurable. Then, by Proposition 3, its pre-image, N, must be measurable as well. But we know it isn't?! $\square$

Let's talk about something more general now, and later return to Proposition 3.

***Definition* 4.** $f : X \to Y$ is continuous, if for every open set $O \subset Y$, $f^{-1}(O) = \{x \in X \mid f(x) \in O\}$ is open in X.

***Definition* 5** (Induced topology)**.** If X is a topological space and $Y \subset X$, the induced topology on Y is defined as follows: $O \subset Y$ is open in Y if there exists an open subset $\tilde{O} \subset X$ in X, such that $O = \tilde{O} \cap Y$.

Theorem 1 (Tietze)**.** If $E \in \mathbb{R}$, E is closed, then any continuous function on E extends to a continuous function on $\mathbb{R}$.

Proof idea. The idea is that if E is closed, then its complement, i.e. the set where f is undefined, is open. As we know, an open set on $\mathbb{R}$ can be represented as a disjoint union of a countable number of intervals. For every such interval we can simply connect the values of f on its endpoints with a straight line.

The technicality that will arise is the case if we have a countable number of intervals that accumulate at a point — that can be left as an exercise. $\square$

***Definition* 6.** Let $(X, \mathcal{A})$, $(Y, \mathcal{D})$ be measurable spaces. $f : X \to Y$ is called $(A, \mathcal{D})$-measurable, if for every $E \in \mathcal{D}$ we have $f^{-1}(E) \in \mathcal{A}$.

Proposition 4. Let $f : X \to Y$ be any function. $\mathcal{A}$ is a σ-algebra on X. Then the family of sets

$$f_*(\mathcal{A}) = \{E \subset Y \mid f^{-1}(E) \in \mathcal{A}\}$$

is a σ-algebra on Y.

Proof. To prove that $f_*(\mathcal{A})$ is a σ-algebra, we have to prove 3 properties from the definition of a σ-algebra.

1. $Y \in f_*(\mathcal{A})$, since $f^{-1}(Y) = X$. And $X \in \mathcal{A}$, since $\mathcal{A}$ is a σ-algebra.

2. $f_*(\mathcal{A})$ is closed under complements. If $E \in f_*(\mathcal{A})$, then $f^{-1}(E^C) = \left(f^{-1}(E)\right)^C$. But $f^{-1}(E) \in \mathcal{A}$ by definition of $f_*(\mathcal{A})$. Therefore, $\left(f^{-1}(E)\right)^C \in \mathcal{A}$, since $\mathcal{A}$ is a σ-algebra. Therefore, $f^{-1}(E^C) \in \mathcal{A}$, and thus $E^C \in f_*(\mathcal{A})$ by definition of $f_*(\mathcal{A})$.

3. $f_*(\mathcal{A})$ is closed under countable unions. If $\{E_n\} \subset f_*(\mathcal{A})$, then

$$f^{-1}\Big(\bigcup_{k=1}^{\infty} E_k\Big) = \bigcup_{k=1}^{\infty} f^{-1}(E_k) \in \mathcal{A} \implies \bigcup_{k=1}^{\infty} E_k \in f_*(\mathcal{A})$$

□

Observation. If $f_*(\mathcal{A})$ contains a generating set of $\mathcal{D}$, then f is $(\mathcal{A}, \mathcal{D})$-measurable.

Proof. If $f_*(\mathcal{A})$ contains a generating set of $\mathcal{D}$, then, since it is a σ-algebra, it must contain the whole $\mathcal{D}$. That means that the preimages of all sets from $\mathcal{D}$ lie in $\mathcal{A}$, which is the definition of $(\mathcal{A}, \mathcal{D})$-measurability. □

Corollary 2. Let $(X, \mathcal{A})$, $(Y, \mathcal{D})$ be measurable spaces. $\mathcal{E} \subset \mathcal{D}$ generates $\mathcal{D}$. Then $f : X \to Y$ is $(\mathcal{A}, \mathcal{D})$-measurable if and only if $\forall E \in \mathcal{E} : f^{-1}(E) \in \mathcal{A}$.

Proof. This is just a restatement of the observation above. □

Corollary 3. Let X, Y be topological spaces. If $f : X \to Y$ is continuous, then f is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable.

Proof. Take $\mathcal{E}$ to be the set of all open sets in Y. $\mathcal{E}$ is a generating set of $\mathcal{B}_Y$. Then, according to the Corollary 2, we have to prove that the inverse of every set in $\mathcal{E}$ is in $\mathcal{B}_X$. But f^{-1} of every open set is open by the definition of a continuous function, and every open set in X is contained in $\mathcal{B}_X$ (since the Borel σ-algebra is by definition the smallest σ-algebra containing all open sets). □

Proof of Proposition 3. Let $X = E$, $Y = \mathbb{R}$ in the previous corollary. If $f : E \to \mathbb{R}$ is continuous, then f is $(\mathcal{B}_E, \mathcal{B}_{\mathbb{R}})$-measurable. But $\mathcal{B}_E \subset \mathcal{M}$. So, the preimage of any set in $\mathcal{B}_{\mathbb{R}}$ will be in $\mathcal{M}$. □

1.8 Measurable functions in $\overline{\mathbb{R}}$

Definition 1. $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$. The topology on $\overline{\mathbb{R}}$ is defined as follow: open sets are all open sets in $\mathbb{R}$, all intervals of the form $(a, +\infty]$, $[-\infty, a)$, as well as all possible unions (not neccesarily countable) of the aforementioned sets.

Definition 2. We can say that a sequence converges to $+\infty$ or $-\infty$, if for any open set containing this infinity there exists such an index of the sequence, starting from which all elements of the sequence are contained in that set.

Definition 3. Let $E \in \mathbb{R}$ be measurable. Then $f : E \to \overline{\mathbb{R}}$ is *(Lebesgue) measurable* if it is $(\mathcal{M}, \mathcal{B}_{\overline{\mathbb{R}}})$-measurable.

Example 1. All continuous functions are Lebesgue measurable.

Example 2. The Dirichlet function (1 for $\mathbb{Q}$, 0 for $\mathbb{R} \setminus \mathbb{Q}$) is Lebesgue measurable. That's because every preimage for the Dirichlet function is either $\mathbb{R}$, $\mathbb{Q}$, or $\mathbb{R} \setminus \mathbb{Q}$, and every one of those is measurable.

Corollary 1. Let $E \in \mathcal{M}$, $f : E \to \overline{\mathbb{R}}$ is measurable if and only if on the of the following holds:

1. For every open set $O \subset \overline{\mathbb{R}} : f^{-1}(O) \in \mathcal{M}$.

2. $\forall c \in \overline{\mathbb{R}} \implies f^{-1}\big((c, +\infty]\big) \in \mathcal{M}$.
3. $\forall c \in \overline{\mathbb{R}} \implies f^{-1}\big([c, +\infty]\big) \in \mathcal{M}$.
4. $\forall c \in \overline{\mathbb{R}} \implies f^{-1}\big([-\infty, c)\big) \in \mathcal{M}$.
5. $\forall c \in \overline{\mathbb{R}} \implies f^{-1}\big([-\infty, c]\big) \in \mathcal{M}$.

Proof. The families of sets described above are generating sets of the Borel σ-algebra. Therefore, the corollary follows directly from this observation. $\square$

Observation. $E \in \mathcal{M}$, $f : E \to \overline{\mathbb{R}}$.

1. If f is measurable and $f = g$ almost everywhere (f and g differ on a subset of measure 0), then g is measurable.
2. Let $D \subset E$, $D \in \mathcal{M}$. f is measurable on $E \iff f|_D$ and $f|_{E \setminus D}$ are measurable.

Proof.
1. Let $B \subset \overline{\mathbb{R}}$. f is measurable, thus, $f^{-1}(B)$ is measurable. f and g differ on a subset of measure 0, therefore, $g^{-1}(B) \triangle f^{-1}(B)$ has measure 0, therefore, $g^{-1}(B)$ is measurable.
2. (a) $\impliedby$
$$f^{-1}(B) = (f|_D)^{-1}(B) \cup (f|_{E \setminus D})^{-1}(B)$$
The union of measurable sets is measurable, therefore, $f^{-1}(B)$ is measurable.

   (b) $\implies$
$$(f|_D)^{-1}(B) = f^{-1}(B) \cap D \text{ and } (f|_{E \setminus D})^{-1}(B) = f^{-1}(B) \cap (E \setminus D)$$
(Intersection of two measurable sets is measurable.)

$\square$

Observation. If $f(x) = +\infty$ and $g(x) = -\infty$, then $(f + g)(x)$ is not defined. Therefore, if f and g are finite almost everywhere, then $f + g$ is defined almost everywhere.

Remark. E is always considered measurable when we write "$f : E \to \mathbb{R}$ is measurable".

Theorem 1. $f, g : E \to \overline{\mathbb{R}}$ — measurable, finite almost everywhere. Then:

1. $\forall \alpha, \beta \in \mathbb{R} : \alpha f + \beta g$ is measurable on E.
2. $f \cdot g$ is measurable on E.

Proof.

1.
   - αf is measurable.
     (a) $\alpha = 0 \implies \alpha f \equiv 0$ — measurable.
     (b) $\alpha \neq 0$. $F = \alpha f \implies F^{-1}(\text{open set}) = f^{-1}(\frac{1}{2} \cdot \text{open set})$. Here $\frac{1}{2} \cdot$ open set is also open, therefore, f^{-1} of it is measurable, therefore, $F^{-1}(\text{open set})$ is measurable.

- $F = f + g$. Consider $F^{-1}([-\infty, c))$ $\forall c \in \mathbb{R}$.

$$F^{-1}([-\infty, c)) = \{x \mid f(x) + g(x) < c\}$$

For any such x there exists a $q \in \mathbb{Q}$, such that

$$f(x) < q < c - g(x) \implies F^{-1}([-\infty, c)) \subset \bigcup_{q \in \mathbb{Q}} \Big(\{x \mid g(x) < c - q\} \cap \{x \mid f(x) < q\}\Big)$$

(Here we have a set intersection: if x satisfies both inequalities, then it is in $F^{-1}([-\infty, c])$.) But

$$\bigcup_{q \in \mathbb{Q}} \Big(\{x \mid g(x) < c - q\} \cap \{x \mid f(x) < q\}\Big) = \bigcup_{q \in \mathbb{Q}} \Big(g^{-1}([-\infty, c - q)) \cap f^{-1}([-\infty, q])\Big)$$

A countable union of measurable sets is measurable, therefore, F is measurable.

2.

- $f \cdot g = \frac{1}{2}\big((f+g)^2 - f^2 - g^2\big)$
- $F = f^2$. For all $c \geqslant 0$, consider

$$F^{-1}\big((c, +\infty]\big) = f^{-1}\big((\sqrt{c}, +\infty]\big) \cup f^{-1}\big([-\infty, -\sqrt{c})\big) \in \mathcal{M}$$

For $c < 0$, $F^{-1}((c, +\infty)) = E$.

$\square$

Example. Composition of measurable functions can be non-measurable: $\psi = \varphi + \mathrm{id}$, φ — Cantor function. Let A — measurable set contained in Cantor set C, such that $\psi(A)$ is not measurable. ψ and ψ^{-1} are continuous, and thus measurable.

$$\chi_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

χ_A is measurable, since A is measurable. But $f = \chi_A \circ \psi^{-1}$ is non-measurable.

Proof.

$$f^{-1}(\{1\} \in \mathcal{B}) = \psi \circ \chi_A^{-1}(\{1\}) = \psi(A) \notin \mathcal{M}$$

Therefore, f is not measurable. $\square$

Proposition 1.

$$\begin{cases} f : \overline{\mathbb{R}} \to \overline{\mathbb{R}} \text{ is continuous} \\ g : E \to \overline{\mathbb{R}} \text{ is measurable} \end{cases} \implies f \circ g \text{ is measurable}$$

Proof. For every open set $O \in \overline{\mathbb{R}}$ we have

$$(f \circ g)^{-1}(O) = g^{-1}\big(f^{-1}(O)\big) \overset{f^{-1}(O) \text{ is open}}{\in} \mathcal{M}$$

$\square$

Corollary 2. For every $p > 0$ and f — measurable function, $|f|^p$ is measurable.

Proof. $|x|$ and x^p are continuous as functions from $\mathbb{R}$ to $\mathbb{R}$. Therefore, $|f|^p$ is continuous as a composition. $\square$

Remark. If $p < 0$, then x^p is not continuous at 0. But we can prove that it is as function from $[0, +\infty]$ to $\overline{\mathbb{R}}$.

***Definition* 4.** $f_1, \dots, f_n : E \to \overline{\mathbb{R}}$. Define

$$\max\{f_1, \dots, f_n\}(x) := \max\{f_1(x), \dots, f_n(x)\}$$

Proposition 2. If $f_1, \dots, f_n : E \to \overline{\mathbb{R}}$ are measurable, then $\max\{f_1, \dots, f_n\}$ is measurable and $\min\{f_1, \dots, f_n\}$ is measurable.

Proof. For all $c \in \mathbb{R}$,

$$\max\{f_1, \dots, f_n\}^{-1}\big((c, +\infty]\big) = \bigcup_{k=1}^{n} f_k^{-1}\big((c, +\infty]\big)$$

which is measurable as a finite union of measurable sets.

(Meaning of the formula: the maximum is greater than c at a point if at least one of the individual functions is greater than c at that point.) $\square$

Remark. In principle, max can be turned to sup or even lim sup.

For supremum, you would consider a countable union, and for lim sup you would consider a countable intersection, both of which don't change the measurability of sets.

***Definition* 5.** For a sequence $\{a_k\}$,

$$\limsup a_n = \lim_{k \to \infty} \sup\{a_k, a_{k+1}, \dots\}$$

Corollary 3. If f is measurable, then $f^+(x) = \max\{f, 0\}$ (called the *positive part* of f) and $f^-(x) = \min\{f, 0\}$ (called the *negative part* of f) are measurable.

Proof. This is the direct sequence of Proposition 2. $\square$

***Definition* 6** (Types of convergence)**.** $f_n : E \to \overline{\mathbb{R}}$.

Then the sequence of functions $\{f_n\}$ converges to f:

1. *pointwise*, if $\lim_{n\to\infty} f_n(x) = f(x), \ \forall x \in E$
2. *pointwise almost everywhere of* E if $\lim_{n\to\infty} f_n(x) = f(x)$ for almost every $x \in E$ (i.e., for every x outside a subset of measure 0).
3. *uniformly*, if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$, such that $|f_n(x) - f(x)| < \varepsilon$ for any $n > N, \ x \in E$.

Remark. Uniform convergence is the strongest one.

Remark. Pointwise limit of continuous functions (or Riemann integrals) can be discontinuous (or not Riemann integrable). For example, if we number all of the rational numbers, and construct continuous functions that raise to the rational numbers and then go down again for the first n rational numbers for every n, then the limit of such functions is the Dirichlet function.

Despite that, the following proposition holds true true:

Proposition 3. $f_n : E \to \overline{\mathbb{R}}$ — measurable. If $f_n \to f$ pointwise almost everywhere on E, then f is measurable.

Proof. Without the loss of generality, the convergence is pointwise. (Because if we prove it for a subset of E of the same measure, then returning the set of points of measure zero to our domain of definition will not change the measurability).

Let's consider $f^{-1}\big([-\infty, c]\big) = \{x \mid f(x) < c\}$.

$$f(x) < c \Longleftrightarrow \exists n, k \in \mathbb{N} : \forall j > k : f_j(x) < c - \frac{1}{n}$$

(Note that this $\frac{1}{n}$ part is important, because if we omitted that, it could happen that $f_j(x)$ converge to c, but are always strictly less than c.) Therefore,

$$f^{-1}\big([-\infty, c)\big) = \bigcup_{k,n} \bigcap_{j=k+1}^{\infty} f_j^{-1}\left(\left[-\infty, c - \frac{1}{n}\right)\right)$$

which is measurable because countable unions and intersections keep measurability. □

1.9 Simple functions

***Definition* 1.** Let $E \in \mathcal{M}$. A function $\varphi : E \to \mathbb{R}$ is called a *simple*, if it is measurable and takes only finitely many values.

Property. A simple function can be represented in the following way (also called the canonical representation):

$$\varphi = \sum_{k=1}^{n} c_k \cdot \chi_{E_k}, \text{ where } E_k = \varphi^{-1}(c_k)$$

(Here c_k are all the values of φ, χ is the characteristic function.)

Remark. We could also represent a simple function in the following way:

$$\varphi = \sum_{k=1}^{n} c_n \chi_{E_k}$$

Here E_k may overlap. But we can easily derive a canonical representation. Say, $E_i \cap E_j \neq \varnothing$. Then

$$\varphi = c_i \cdot \chi_{E_i \setminus E_j} + c_j \cdot \chi_{E_j \setminus E_i} + (c_i + c_j) \cdot \chi_{E_i \cap E_j} + \dots$$

Lemma (Simple approximation lemma)**.** $f : E \to \mathbb{R}$ — measurable and bounded ($\exists M : |f| < m$). Then for every $\varepsilon > 0$ there exist simple functions $\varphi_\varepsilon, \psi_\varepsilon : E \to \mathbb{R}$, such that $\varphi_\varepsilon \leqslant f \leqslant \psi_\varepsilon$ and $0 \leqslant \psi_\varepsilon - \varphi_\varepsilon < \varepsilon$ on E.

Proof. Let's split the range of function f into equal intervals $I_1, \dots, I_n$ of length no more than ε. Now, for every $x \in E$, let's round $f(x)$ *down* to the closest interval end point and call the resulting function $\varphi_\varepsilon(x)$, and let's round $f(x)$ *up* to the closest interval end point and call the resulting function $\psi_\varepsilon(x)$.

Then indeed we'll have $\varphi_\varepsilon \leqslant f \leqslant \psi_\varepsilon$ (since we rounded the function down or up, respectively), and $0 \leqslant \psi_\varepsilon - \varphi_\varepsilon < \varepsilon$ as the rounded points will end in the same interval for every $x \in E$. □

Theorem 1 (Simple approximation theorem)**.** the:simpleApproxTheorem $f : E \to \overline{\mathbb{R}}$ is measurable if and only if there exists a sequence of simple functions $\varphi_n : E \to \mathbb{R}$, such that $\varphi_n \to f$ pointwise on E, and $|\varphi_n| \leqslant |f|$ on E for all n.

(If $f \geqslant 0$, then $\{\varphi_n\}$ can be chosen to be increasing.)

Remark. In the simple approximation lemma, functions φ_ε and ψ_ε sandwich the function f, and $0 \leqslant \psi_\varepsilon - \varphi_\varepsilon < \varepsilon$, therefore, at every point they both differ from f by no more than ε. If we could choose a sequence of ε's that converges to 0, we would get uniform convergence.

Unfortunately, we can't apply the lemma directly, because we needed the function f to be bounded. That's why we're not getting uniform convergence.

Proof. 1. $\Longleftarrow$. We have a sequence of simple (and thus measurable) functions $\{\varphi_n\}$, therefore, their limit is also measurable.

2. $\Longrightarrow$.

   (1) Let's express $f = f^+ - f^-$, where $f^+ = \max\{f, 0\}$ and $f^- = \min\{f, 0\}$. Let's prove the theorem for both f^+ and f^-. Without the loss of generality, assume that $f \geqslant 0$.

   (2) To use the simple approximation lemma, we have to make the function bounded. Define $f_n \coloneqq \min\{f, 2^n\}$. Take φ_n as the lower simple approximation of f_n, according to the simple approximation lemma, with $\varepsilon = \frac{1}{2^n}$.
   Then $\varphi_n \to f$ pointwise. Moreover, $\{\varphi_n\}$ are increasing (because we bound it with 2^n, and because when we go from n to $n+1$, every interval in the proof of SAL is split into two).

   (3) Do the same for f^- and add the sequences for f^+ and f^- together. Since $|f| = \max\{|f^+|, |f^-|\}$, the condition of $|\varphi_n| \leqslant |f|$ will hold true.

□

1.10 Egoroff, Lusin, Littlewood's three principles

Informal statement:

1. Every measurable set is *nearly* (not a mathematical term) a finite union of intervals.
2. Every measurable function is nearly continuous.
3. Every pointwise convergent sequence of measurable functions is nearly uniform convergent.

Theorem 1 (Egoroff's theorem)**.** Assume that $m(E) < \infty$, $f_n : E \to \overline{\mathbb{R}}$ — measurable, $f_n \to f$ pointwise almost everywhere on E, where $f : E \to \mathbb{R}$. Then for any $\varepsilon > 0$ there exists a closed subset $F \subset E$, such $m(E \setminus F) < \varepsilon$ and $f_n \to f$ uniformly on f.

Remark. It is crucial that f is from E to $\mathbb{R}$, not $\overline{\mathbb{R}}$, otherwise we wouldn't get uniform convergence.

Proof. $|f| < \infty \implies |f - f_k|$ is defined for all $x \in E$ (starting from some k).

Let's also throw out the subset of measure E, on which $f_n \not\to f$.

1. Fix $\eta > 0$.
$$E_n \coloneqq \{x \in E \mid |f(x) - f_k(x)| < \eta,\ \forall k \geqslant n\}$$
Each set E_n is measurable, because $|f(x) - f_k(x)|$ is measurable, and E_n is the countable intersection of preimages of $[0, \eta)$ for all k. We can also note that $\{E_n\}$ is ascending.
Also, $\cup_{n=1}^{\infty} E_n = E$. That's follows from pointwise convergence. If we take any point from the set E, eventually it will be contained in some E_n, because for a given point x the difference $|f(x) - f_k(x)|$ has to converge to 0. Therefore,
$$\lim_{n\to\infty} m(E_n) = m(E)$$
Then there exists $N \in \mathbb{N}$, such that $m(E \setminus E_n) < \varepsilon$ for every $\varepsilon > 0$.

2. Now let's take such N for $\eta_n = \frac{1}{n}$, $\varepsilon_n = \frac{\varepsilon}{2^{n+1}}$.

   Take $E_n = E_N$, where E_N is from step 1 for $\eta = \eta_n$ and $\varepsilon = \varepsilon_n$.

   Take $A = \cap_{n=1}^{\infty} E_n$. Then $f_n \to f$ uniformly on A. On the other hand, A is measurable as a countable union of measurable sets.

   Let's estimate the measure of A.

$$m(E \setminus A) = m\Big(\bigcup_{k=1}^{\infty}(E \setminus E_n)\Big) \leqslant \sum_{n=1}^{\infty} m(E \setminus E_n) \leqslant \sum_{n=1}^{\infty} \frac{\varepsilon}{2^{n+1}} = \frac{\varepsilon}{2}$$

   But A may be non-closed, so we can't just assign F to A. Therefore, we have to approximate A by closed sets.

3. Take $F \subset A$, so that F is closed and $m(A \subset F) < \frac{\varepsilon}{2}$. We have proved that such an F exists if A is measurable.

□

Theorem 2 (Lusin's theorem)**.** $f : E \to \mathbb{R}$ — measurable. Then for every $\varepsilon > 0$ there exists a continuous function $g : \mathbb{R} \to \mathbb{R}$ and a closed set $F \subset E$, such that $f = g$ on F and $m(E \setminus F) < \varepsilon$.

Proof. *Main idea*: uniform limit of continuous functions is continuous, then apply Egoroff's theorem.

The proof is gonna be split into two parts:

1. Find a closed F on which f is continuous (in induced topology).

2. Extend $f|_F$ to a continuous function g. We have briefly discussed how to do this: the complement of a closed set F is an open set, and thus a disjoint union of open intervals. On every such open interval we can connect the graph of f with a straight line. There's some work to be done, because there can be an infinite number of such intervals converging a point, but we'll leave it as an exercise.

Proof of Part 1:

1. Assume $m(E) < \infty$. (Let's leave $m(E) = \infty$ as an exercise). By simple approximation lemma, there exists a sequence of simple functions $f_n : E \to \mathbb{R}$, such that $f_n \to f$ pointwise.

2. $\forall n$ there exists a closed subset $F_n \subset F$, such that f_n is continuous on F_n and $m(E \setminus F_n) < \frac{\varepsilon}{2^{n+1}}$.

   Proof:

$$f_n = \sum_{j=1}^{k_n} c_j \cdot \chi_{E_j} \text{ — canonical representation}$$

   Each of the sets E_j can be approximated by a closed set G_j from the inside (they are measurable from the definition of a simple function).

   Put $F_n \coloneqq \cup G_j$. It's a finite union, therefore F_n is closed.

   f_n is constant on each of G_j, as $G_j \subset E_j$. A preimage of an open set for f was a union of several E_j's. If we now restrict f onto $\cup G_j$, the preimage of an open set will be a union of several G_j's. What we have to prove that G_j's are open in the induced topogy on F_n — that's true because G_j's are closed in $\mathbb{R}$ and disjoint.

3. By Egoroff's theorem, there exists a closed $F_0 \subset E$, such that $m(E \setminus F_0) < \frac{\varepsilon}{2}$ and $f_n \to f$ uniformly. Now take $F := \bigcap_{n=0}^{\infty} F_n$. The intersection of closed sets is closed. By construction, f_n is continuous on F, as $F_n \subset F$. Since $f_n \to f$ uniformly on F (as $F_0 \subset F$), f is continuous on F. Let's now estimate the measure of $E \setminus F$:

$$m(E \setminus F) < \frac{\varepsilon}{2} + \sum_{n=1}^{\infty} \frac{\varepsilon}{2^{n+1}} = \varepsilon$$

□

Remark. If $m(E) = \infty$, we can split $\mathbb{R}$ into intervals with finite length. Then use the finite case we've just proved for each of those intervals, but choose $\varepsilon_k = \frac{\varepsilon}{2^{k+1}}$ so that they all sum up to ε at the end.

2 Lebesgue integral

***Definition* 1.** Let $\psi : E \to \mathbb{R}$ be simple, $m(E) < \infty$. Let

$$\psi = \sum_{n=1}^{k} a_n \cdot \chi_{E_n}$$

be the canonical representation of ψ. Then define

$$\int_E \psi := \sum_{n=1}^{k} a_n \cdot m(E_n)$$

Remark. Notation: if we want to emphasize the measure, we can write

$$\int_E \psi = \int_E \psi \, dm = \int_a^b \psi \, dm \text{ (when } E = [a, b])$$

Remark. One can show that if it's not the canonical representation, we can still compute the integral by the same formula and get the same result.

***Definition* 2.** Let $f : E \to \mathbb{R}$ be bounded, $m(E) < \infty$. (f is not necessarily measurable).

The *lower Lebesgue integral* is by definition

$$\sup\left\{ \int_E \varphi \;\middle|\; \varphi \text{ is simple and } \varphi \leqslant f \text{ on } E \right\}$$

The *upper Lebesgue integral* is by definition

$$\inf\left\{ \int_E \psi \;\middle|\; \varphi \text{ is simple and } \varphi \geqslant f \text{ on } E \right\}$$

f is *Lebesgue integrable* if the lower Lebesgue integral is equal to the upper Lebesgue integral. Then

$$\int_E f := \text{lower Lebesgue integral} = \text{upper Lebesgue integral}$$

Theorem 1. $f : [a, b] \to \mathbb{R}$ is bounded and $[a, b]$ is bounded. If f is Riemann integrable over $[a, b]$, then f is Lebesgue integrable and both integrals will match.

Proof. Upper and lower Riemann sums are obtained by simple functions. □

Theorem 2. If $f : E \to \mathbb{R}$ is bounded and *measurable*, $m(E) < \infty$, then f is Lebesgue integrable on E.

Proof. By the simple approximation lemma, for every $n \in \mathbb{N}$ there exist simple functions φ_n, ϕ_n, such that $\varphi_n \leqslant f \leqslant \psi_n$ on E and $\psi_n - \varphi_n \leqslant \frac{1}{n}$ on E.

$$\forall n : (\text{upper Lebesgue integral} - \text{lower Lebesgue integral}) \leqslant \int_E \psi_n - \int_E \varphi_n = \int_E (\psi_n - \varphi_n) \leqslant \frac{1}{n} m(E) \to 0$$

Here we've used the fact that the difference of two integrals is the integral of the difference (which we haven't yet proved). However, since we're dealing with simple functions, we can prove it by simply rearranging the finite sums. □

Now here's some definitions for the improper case, i.e. infinite integration limits or functions that converge to infinity.

Definition. $f : E \to \overline{\mathbb{R}}$ is measurable. Its support is by definition

$$\operatorname{supp} f \coloneqq \{x \in E \mid f(x) \neq 0\}$$

f is of finite support of $m(\operatorname{supp} f) < \infty$.

Definition 3. $f : E \to \overline{\mathbb{R}}$ is measurable, $f \geqslant 0$. Then

$$\int_E f \coloneqq \sup\Big\{ \int_E h \;\Big|\; h \text{ is bounded, measurable, } m(\operatorname{supp} h) < \infty,\ 0 \leqslant h \leqslant f \text{ on } E \Big\}$$

f is Lebesgue integrable over E, if $\int_E f < \infty$.

Definition 4. A measurable function $f : E \to \overline{\mathbb{R}}$ is Lebesgue measurable over E if $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$ are Lebesgue integrable over E. Then

$$\int_E f = \int_E f^+ - \int_E f^-$$

Definition 5. A function $f : E \to \overline{\mathbb{R}}$ is integrable if both $\int_E f^+$ and $\int_E f^-$ are finite.

Remark. If $\int_E f^+$ and $\int_E f^-$ are both infinite, we can't subtract them in the previous definition. Therefore, it is not enough to just say that $\int_E f$ has to be finite.

Theorem 3. $f, g : E \to \mathbb{R}$ are measurable and integrable. Then for any $\alpha, \beta \in \mathbb{R}$ the linear combination $\alpha f + \beta g$ is integrable, and

$$\int_E (\alpha f + \beta g) = \alpha \int_E f + \beta \int_E g$$

Theorem 4. $f, g : E \to \mathbb{R}$ are measurable and non-negative. Then for any $\alpha, \beta \geqslant 0$:

$$\int_E (\alpha f + \beta g) = \alpha \int_E f + \beta \int_E g$$

Proof. The proof of these two theorems is technical, we are not gonna prove them here. You can read them in the textbook. Checking the linearity of simple functions is easy. But the moment we proceed to supremums and infimums it becomes more difficult. □

Remark. Why do we need two theorems instead of one? Notice that Theorem 4 doesn't require integrability. If f and g are non-negative, then their integrals are non-negative as well, therefore, we can add the two integrals and not worry about cases like $+\infty + (-\infty)$.

In Theorem 3, we require integrability. Therefore, the integrals of f and g are both finite, therefore, the right-hand side is well-defined. But in $\int_E (\alpha f + \beta g)$ we could indeed have $+\infty + (-\infty)$. But that's not a problem: we will prove later, that if a function is integrable, then it's infinite only on a subset of measure 0.

2.1 Basic properties of Lebesgue integral

Theorem 1 (Monotonicity)**.** Assume $f, g : E \to \overline{\mathbb{R}}$ are measurable and either both integrable or both non-negative. Then if $f \leqslant g$ on E, then

$$\int_E f \leqslant \int_E g$$

Proof. We can use either Theorem 3 or Theorem 4. If $g = f + h$, then $g \geqslant f \implies h \geqslant 0 \implies \int_E h \geqslant 0$. But $\int_E g = \int_E f + \int_E h$ by one of those theorems, therefore, $\int_E g \geqslant \int_E f$. □

Proposition 1. $f : E \to \overline{\mathbb{R}}$ is measurable. Then f is integrable $\iff$ $|f|$ is integrable.

Proof. $|f| = f^+ + f^-$.

1. $\implies$. If f is integrable, then $\int f^+$, $\int f^- < \infty$ by definition. Then
$$\int |f| = \int f^+ + \int f^- < \infty$$
2. $\impliedby$. $0 \leqslant f^+ \leqslant |f|$ and $0 \leqslant f^- \leqslant |f|$. If $|f|$ is integrable, then, by monotonicity, f^+ and f^- are integrable.

□

Proposition 2. $f : E \to \overline{\mathbb{R}}$, $g : E \to \overline{\mathbb{R}}$ — integrable and $g \geqslant 0$. If $|f| \leqslant g$ on E, then f is integrable.

Proof. If g is integrable, then by monotonicity $|f|$ is integrable, then by previous proposition f is integrable. □

Theorem 2. $f : E \to \overline{\mathbb{R}}$ is measurable and either integrable or non-negative. If $A, B \subset E$, $A \cap B = \varnothing$; $A, B \in \mathcal{M}$, then
$$\int_{A\cup B} f = \int_A f + \int_B f$$

Proof.
$$f\big|_{A\cup B} = f \cdot \chi_A + f \cdot \chi_B = h + g$$
If f is integrable, then h and g are integrable by previous proposition, because $|h| \leqslant |f|$ and $|g| \leqslant |f|$. Now apply Theorem 3.

If f is non-negative, then $h, g \geqslant 0$, apply Theorem 4. □

Proposition 3. If $f : E \to \overline{\mathbb{R}}$ is integrable, then
$$\left|\int_E f\right| \leqslant \int_E |f|$$

Proof.
$$\left|\int_E f\right| = \left|\int_E f^+ - \int_E f^-\right| \leqslant \int_E f^+ + \int_E f^- = \int_E |f|$$

□

Theorem 3 (Chebyshev's inequality)**.** $f : E \to \overline{\mathbb{R}}$ — measurable, $f \geqslant 0$. Then $\forall \lambda > 0$:
$$m(E_\lambda) = m\{x \in E \mid f(x) \geqslant \lambda\} \leqslant \frac{1}{\lambda}\int_E f$$

Proof. Put $h \coloneqq \lambda \cdot \chi_{E_\lambda}$. Since h is a simple function,
$$\int_{E_\lambda} h = \lambda \cdot m(E_\lambda)$$
Therefore, it is enough to prove that
$$\int_{E_\lambda} h \leqslant \int_E f$$
But $h \leqslant f\big|_{E_\lambda}$, because on E_λ we have $h = \lambda$ and $f \geqslant \lambda$. Therefore, by monotonicity,
$$\int_{E_\lambda} h \leqslant \int_{E_\lambda} f \leqslant \int_E f$$

□

Applications of the Chebyshev's inequality:

Proposition 4. $f : E \to \overline{\mathbb{R}}$ is measurable, $f \geqslant 0$. Then

$$\int_E f = 0 \iff f = 0 \text{ almost everywhere on } \mathbb{R}$$

Proof.

1. $\implies$.
   If $\int_E f = 0$, then by Chebyshev's inequality,
   $$m(E_n) = m\Big\{x \in E \;\Big|\; f(x) \geqslant \frac{1}{n}\Big\} \leqslant n \int_E f = 0$$
   The sequence $\{E_n\}$ is ascending. Since the measure of each one of them is equal to zero, then, by continuity of measure, $m(\cup_{n=1}^{\infty} E_n) = m\{x \in E \mid f(x) > 0\} = 0$.
2. $\impliedby$. Follows from Definition 3.

□

Proposition 5. If $f : E \to \overline{\mathbb{R}}$ is integrable, then f is finite almost everywhere on E.

Proof. Without the loss of generality, assume that $f \geqslant 0$ (otherwise, prove separately for f^+ and f^-).

$$m\{x \in E \mid f(x) = \infty\} \leqslant m\{x \in E \mid f(x) \geqslant n\} \overset{\text{Chebyshev's inequality}}{\leqslant} \frac{1}{n} \int_E f \underset{n\to\infty}{\longrightarrow} 0$$

Here we've used that $\int_E f$ is finite, because f is integrable. □

2.2 Convergence theorems

Proposition 1. $f_n : E \to \overline{\mathbb{R}}$ — measurable, integrable, $m(E) < \infty$. If $f_n \to f$ uniformly and f is bounded on E, then

$$\lim_{n\to\infty} \int_E f_n = \int_E f$$

Proof. f is measurable (as a pointwise limit). Therefore, $\int_E f$ exists. For any $\varepsilon > 0$, choose $N > 0$, such that

$$|f - f_n| < \frac{\varepsilon}{m(E)}, \quad \forall n \geqslant N$$

Then

$$\left|\int_E f - \int_E f_n\right| = \left|\int_E (f - f_n)\right| \leqslant \int_E |f - f_n| < \frac{\varepsilon}{m(E)} \cdot m(E) = \varepsilon$$

□

Remark. This proposition does not hold for pointwise convergence.

Example 1. $E = (0, 1]$, $f_n := n \cdot \chi_{(0,1/n)}$. Then $f_n \to 0$ pointwise. But:

$$\int_E f_n = 1 \quad \int_E f = 0$$

Example 2. $E = \mathbb{R}$, $f_n = \chi_{[n,n+1]}$. If we choose an x, then for any $n > x$ this interval will be to the right of x, and thus $f_n(x) = 0$. Therefore, $f_n \to 0$ pointwise. But:

$$\int_E f_n = 1 \quad \int_E f = 0$$

Theorem 1 (Bounded convergence theorem)**.** $f_n : E \to \overline{\mathbb{R}}$ — measurable, $m(E) < \infty$. Assume that there exists an $M > 0$, such that $|f_n| < M$ on E for all n. Then if $f_n \to f$ pointwise on E, then

$$\lim_{n\to\infty} \int_E f_n = \int_E f$$

Remark. If we look at the two examples stated earlier as a movie about a function, with frames f_n, then we can see how the "mass" of a function *(not formal)* is moving. In the first example, the "mass" of the function was escaping vertically. Here it's prevented by $|f_n| < M$. In the second example, the "mass" was escaping horizontally. Here it's prevented by $m(E) < \infty$.

Proof. We're gonna use Egoroff's theorem.

1. f is measurable (as a pointwise limit). $|f| \leqslant M$.

2. For any $\varepsilon > 0$, take $A \subset E$, such that $m(E \setminus A) < \frac{\varepsilon}{4M}$ and $f_n \to f$ uniformly on A (by Egoroff's theorem). Then,

$$\left| \int_E f_n - \int_E f \right| \leqslant \left| \int_A f_n - \int_A f \right| + \left| \int_{E\setminus A} f_n - \int_{E\setminus A} f \right| = (*)$$

Since $f_n \to f$ uniformly on A, we have (by Proposition 1) that $\left| \int_A f_n - \int_A f \right| \to 0$ as $n \to \infty$. Now,

$$\left| \int_{E\setminus A} f_n - \int_{E\setminus A} f \right| \leqslant 2M \cdot m(E \setminus A) < \frac{\varepsilon}{2}$$

Since we can make the first modulus be smaller than $\frac{\varepsilon}{2}$ by choosing a sufficiently large n, we get that

$$(*) \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

□

Lemma (Fatou's lemma)**.** $f_n : E \to \overline{\mathbb{R}}$ — measurable, $f_n \geqslant 0$. If $f_n \to f$ pointwise almost everywhere on E, then

$$\int f \leqslant \liminf_{n\to\infty} \int_E f_n$$

Remark. The sequence of integrals may not have a limit, therefore, we have to use $\liminf$ in the statement. By definition,

$$\liminf_{n\to\infty} a_n = \lim_{k\to\infty} \inf_{n\geqslant k} a_n$$

The since $\inf_{n\geqslant k} a_n$ is a non-decreasing sequence, the limit (possibly infinite) always exists.

Remark. If we think in terms of our escape of the mass analogy, the theorem states that mass can escape, but can't enter the limit.

Proof. Exclude a set of measure zero (by Egoroff's theorem), now the convergence is pointwise. $f \geqslant 0$ and f is measurable.

Recall Definition 3. Take $h : E \to \overline{\mathbb{R}}$ to be any bounded measurable function with $m(\operatorname{supp} h) < \infty$ and $0 \leqslant h \leqslant f$ on E, like in that definition.

Let M be such that $h \leqslant M$ on E (since h is bounded) and $E_0 := \operatorname{supp} h$. Define $h_n : E \to \mathbb{R}$ by $h_n := \min\{h, f_n\}$. Then:

1. h_n is measurable, because both h and f_n are measurable by their respective definitions, and h_n is the minimum of two measurable functions.

2. $0 \leqslant h_n \leqslant M$, because $h \leqslant M$ on E.

3. $\operatorname{supp} h_n \subset E_0$, as $h_n \leqslant h$.

4. $h_n \to h$ pointwise on E. That's because f_n converges to f pointwise and $h \leqslant f$. Therefore, at every point, starting from some n, Thus, h will be either less than f_n or arbitrarily close to it, if $h = f$ at that point. Since $h_n = \min\{f, h\}$ and $\min\{f, h\} \to h$, we get that $h_n \to h$.

By Bounded convergence theorem:

$$\lim_{n\to\infty} \int_E h_n \overset{\operatorname{supp} h_n \subset E_0}{=} \lim_{n\to\infty} \int_{E_0} h_n \overset{\text{BCT}}{=} \int_{E_0} h \overset{\operatorname{supp} h = E_0}{=} \int_E h$$

On the other hand,

$$h_n \leqslant f_n \implies \int_E f_n \geqslant \int_E h_n \implies \int_E h \leqslant \liminf_{n\to\infty} \int_E f_n$$

Since h is arbitrary, by Definition 3,

$$\int_E f \leqslant \liminf_{n\to\infty} \int_E f_n$$

□

Theorem 2 (Monotone convergence theorem)**.** $f_n : E \to \overline{\mathbb{R}}$, $f_n \geqslant 0$, measurable. The sequence $\{f_n\}$ is increasing on E ($f_n \geqslant f_k$ for $n > k$). If $f_n \to f$ pointwise almost everywhere on E, then

$$\lim_{n\to\infty} \int_E f_n = \int_E f$$

Proof.

1. From Fatou's lemma,
$$\int_E f \leqslant \liminf \int_E f_n$$

2. $$f_n \nearrow \implies \int_E f \geqslant \limsup \int_E f_n \implies \text{all three are equal}$$

□

Theorem 3 (Lebesgue's dominated convergence theorem)**.** Let $f_n : E \to \overline{\mathbb{R}}$ — measurable. Let $g : E \to \overline{\mathbb{R}}$ be integrable over E, $g \geqslant 0$. Assume that $|f_n| \leqslant g$ for all n.

If $f_n \to f$ pointwise almost everywhere on E, then f is integrable on E and

$$\lim_{n\to\infty} \int_E f_n = \int_E f$$

Remark. Bounded convergence theorem is a special case with

$$g(x) = \begin{cases} M, & x \in E \\ 0, & x \notin E \end{cases}$$

Proof. The idea of the proof is to apply Fatou's lemma to $g + f$ and $g - f$.

1. Since $|f_n| \leqslant g$, we have $|f| \leqslant g$ almost everywhere on E. Therefore, f is integrable as we proved earlier.
2. Remove a subset of measure 0 so that f and each f_n are finite on E. That's possible, because each of the functions f and f_n has an infinite value on a subset of measure 0, and the countable union of subsets of measure 0 also has measure 0. Now $g \pm f$ and $g \pm f_n$ are well-defined.
3. Apply Fatou's lemma:
   (a) $f_n \to f$, and $g - f_n$ are non-negative due to the dominance by g.

   $$\int_E (g-f) \leqslant \liminf \int_E (g - f_n) = \int_E g - \limsup \int_E f_n$$

   Now we can cancel out $\int_E g$ on both sides, since it's finite. We get that

   $$\int_E f \geqslant \limsup \int_E f_N$$

   (b) Apply Fatou's lemma to the sequence of $g + f_n$, which is also non-negative due to the dominance by g:

   $$\int_E (g+f) \leqslant \liminf \int_E (g + f_n) = \int_E g + \liminf \int_E f_n \implies \int_E f \leqslant \liminf \int_E f_n$$

   $$\text{(a)} + \text{(b)} \implies \int_E f = \lim_{n\to\infty} \int_E f_n$$

□

Theorem 4 (Countable additivity of integration)**.**
$f : E \to \overline{\mathbb{R}}$ — integrable. $\{E_n\}^\infty$ are disjoint measurable subsets of E, $E = \cup_1^\infty E_n$. Then

$$\int_E f = \sum_{n=1}^{\infty} \int_{E_n} f$$

Proof. Let $\chi_n := \chi_{\cup_1^n E_k}$, $f_n := f \cdot \chi_n$. Then f_n are measurable and $|f_n| \leqslant |f|$ on E and $f_n \to f$ pointwise on E. Hence by Lebesgue's dominance convergence theorem (with $g = f$),

$$\int_E f = \lim_{n\to\infty} \int_E f_n \overset{(*)}{=} \sum_1^\infty \int_{E_n} f$$

Here (*) holds true, because the for a finite sum this equality is true by applying induction to this theorem, and then we can take the limits. □

Theorem 5 (Continuity of integration)**.** $f : E \to \overline{\mathbb{R}}$ — integrable.

1. If $\{E_n\}_1^\infty$ — ascending, each E_n is measurable, then

$$\int\limits_{\cup_1^\infty E_n} f = \lim_{n\to\infty} \int_{E_n} f$$

2. If $\{E_n\}_1^\infty$ — descending, each E_n is measurable, then

$$\int\limits_{\cap_1^\infty E_n} f = \lim_{n\to\infty} \int_{E_n} f$$

Proof. We're not gonna give the full proof, just the proof idea. We can use the dominated convergence theorem.

In the first case, define $f_n \coloneqq f \cdot \chi_{E_n}$. Since it's an ascending sequence, we'll have $f_n \to f$ pointwise on $\cup_1^\infty E_n$. Now apply dominated convergence theorem with $g = f$ on $\cup_1^\infty E_n$.

The second case is analogous, but we'll have to look at the complement. □

3 L^p-spaces

***Definition* 1.** E — measurable, $1 \leqslant p < \infty$. $\hat{L}^p(E)$ is the collection of measurable functions $f : E \to \overline{\mathbb{R}}$, such that $|f|^p$ is integrable over E.

Remark. This is the definition $\hat{L}^p(E)$. The definition of $L^p(E)$ is almost the same except a small detail, we will fix it later.

Remark. An integrable function has a *finite* integral.

Properties of L^p:

***Definition* 2** (Completeness of reals)**.** If $\lim_{n,m\to\infty}|a_n - a_m| = 0$, then there exists $a \in \mathbb{R}$, such that $\lim_{n\to\infty}|a_n - a| = 0$.

Remark. An equivalent formulation: if $\{a_n\}_1^\infty$ is a Cauchy sequence, then there exists $a \in \mathbb{R}$, such that $\lim_{n\to\infty} a_n = a$.

- Completeness of $L^p(E)$: if
$$\lim_{n,m\to\infty} \int |f_n - f_m|^p = 0$$
then there exists $f \in L^p(E)$, such that
$$\lim_{n\to\infty} \int_E |f_n - f|^p = 0$$
This is Riesz–Fischer theorem.

- L^p is a normed linear space with completeness. With completeness, it follows that L^p is a Banach space.

- L^p is separable: there exists a countable dense subset in L^p.

3.1 Normed linear spaces

We are going to only consider linear spaces over $\mathbb{R}$.

Examples:

1. $C[a,b]$ — the space of all continuous functions on $[a,b]$. The sum of two continuous functions is continuous. The null vector is $f(x) \equiv 0$. We'll leave the full proof as an exercise.

2. $B[a,b]$ — the space of all bounded functions on $[a,b]$. If we take the sum of two bounded functions, it will be bounded as well. Multiplying by a scalar doesn't break the boundedness, either.

3. $\hat{L}^p(E)$ — the space of f, such that $|f|^p$ is integrable. We need to prove that $f + g \in \hat{L}^p(E)$ for $f, g \in \hat{L}^p(E)$, i.e. that $|f+g|^p$ is integrable.

   Proof. f and g are infinite on a subset of measure 0, therefore, $f+g$ is undefined on a subset of measure 0. Let's consider the case if both f and g are finite:
$$|f+g|^p \leqslant \big(2\max(|f|,|g|)\big)^p \leqslant 2^p\big(|f|^p + |g|^p\big)$$
   $|f|^p$ and $|g|^p$ are both integrable, therefore, their sum is integrable, and if we multiply it by a constant (2^p) it's still integrable. □

***Definition* 1** (Norm)**.** X — linear space. $\|\cdot\|$ — a real-valued function on X, such that

1. $\|f+g\| \leqslant \|f\| + \|g\|$
2. $\|\alpha f\| = |\alpha| \cdot \|f\|$.
3. $\|f\| \geqslant 0$ and $|f| = 0$ if and only if $f = \vec{0}$.

Remark. A norm defines a metric: $l(f,g) \coloneqq |f-g|$. However, it doesn't work in the other way: we wouldn't get property 3 from a metric.

Examples:

1. $X = C[a,b],\ \|f\| \coloneqq \max_{x\in[a,b]}|f(x)|$.
2. $X = B[a,b],\ \|f\| \coloneqq \sup_{x\in[a,b]}|f(x)|$.
3. $$X = \hat{L}^p(E),\ \|f\|_p \coloneqq \left(\int_E |f|^p\right)^{\frac{1}{p}}$$

Here we take the root of power p, because property 2 (multiplication by scalar) has to hold.

However, property 3 doesn't hold, because if f is 0 almost everywhere, then $\|f\| = 0$, but $f \neq 0$. We can fix this by putting functions that differ on a subset of measure zero into a single equivalence class: $f \sim g$ if $f = g$ almost everywhere. Now we arrive at the correct definition of L^p:

***Definition* 2.** $L^p(E) \coloneqq \hat{L}^p(E)/\sim$

Remark. If $X(f) = 0$ for an $f \in \hat{L}^p(E)$, then f is non-zero on a subset of measure 0, therefore, it's in the same equivalence class as $g \equiv 0$.

Remark. When we write $f \in L^p(E)$, we will imply that we're talking about the corresponding equivalence class $[f]$.

***Definition* 3.** $f : E \to \overline{\mathbb{R}}$ is *essentially bounded*, if there exists $M \geqslant 0$, such that $f(x) \leqslant M$ for almost every $x \in E$. M is called an *essential upper bound*.

***Definition* 4.** $L^\infty(E)$ is the collection of all equivalence classes $[f]$ of measurable functions f that are essentially bounded.

$\|f\|_\infty$ is the infimum of all essential upper bounds of f. It is also called the essential supremum.

3.2 Useful inequalities

***Definition* 1** (Conjugate)**.** For any $p \in (1,+\infty)$, we will call $q \in (1,+\infty)$ its *conjugate*, if

$$\frac{1}{p} + \frac{1}{q} = 1$$

We will denote it as $q = \overline{p}$.

Remark. This is not the same conjugate as in complex analysis.

Remark. It's natural to say that the conjugate of 1 is $+\infty$, and vice versa.

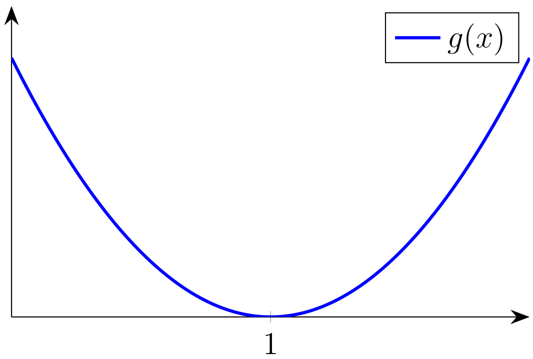
Theorem 1 (Young's inequality)**.** For all $a, b > 0,\ 1 < p < \infty$ and $q = \overline{p}$, we have

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}$$

Proof. Consider $g(x)$:

$$\begin{aligned}
g(x) &= \frac{x^p}{p} + \frac{1}{q} - x \\
g(1) &= \frac{1}{p} + \frac{1}{q} - 1 = 0 \\
g'(x) &= x^{p-1} - 1 \implies g'(x) > 0 \text{ for } x > 1 \text{ and } g'(x) < 0 \text{ for } x \in (0,1)
\end{aligned}$$

So, the derivative to the right of 1 is positive, and the derivative to the left of 1 is negative.



Therefore, $g(1) = 0$ is the minimum of g, thus:

$$\begin{aligned}
& g(x) \geqslant 0 \quad \forall x \in (0, +\infty) \implies \frac{x^p}{p} + \frac{1}{q} \geqslant x \\
& \text{Take } x = \frac{a}{b^{q-1}} \implies \frac{a^p}{pb^{p(q-1)}} + \frac{1}{q} \geqslant \frac{a}{b^{q-1}} \implies \frac{a^p}{b^{pq-p-q}} + \frac{b^q}{q} \geqslant ab \\
& \frac{1}{p} + \frac{1}{q} = 1 \overset{\times pq}{\implies} p + q = pq \implies pq - p - q = 0 \\
& \implies \frac{a^p}{pb^0} + \frac{b^q}{q} \geqslant ab \implies \frac{a^p}{p} + \frac{b^q}{q} \geqslant ab
\end{aligned}$$

□

Remark. This function g looks a bit artificial. It's not the only choice of such function, but it's not the key idea of the proof. The main idea is to create an inequality using derivatives (as we did).

Theorem 2. E — measurable, $1 \leqslant p < \infty$, $q = \overline{p}$. If $f \in L^p(E)$ and $g \in L^q(E)$, then:

1. $f \cdot g \in L^1(E)$ and
$$\int_E |fg| \leqslant \|f\|_p \cdot \|g\|_q \qquad \text{(Hölder's inequality)}$$

2. If f is non-zero on a subset of positive measure (i.e. $[f] \neq [0]$), then
$$\begin{aligned}
& f^*(x) := \|f\|_p^{1-p} \cdot \operatorname{sgn}(f(x)) \cdot |f(x)|^{p-1} \in L^q(E),\ \|f^*\|_q = 1 \\
& \text{and } \int_E f \cdot f^* = \|f\|_p
\end{aligned}$$

Remark. f^* is, in a sense, a rescaling of $|f(x)^{p-1}|$. We multiply it by the sign function to preserve the sign of f.

Proof.

Case 1. $p = 1$, $q = \infty$. In this case Hölder's inequality is obvious if we replace g with its essential supremum:

$$\int_E |fg| \leqslant \int_E |f| \cdot \operatorname{ess\,sup}_E |g| = \int_E |f| \cdot \|g\|_q = \|f\|_p \cdot \|g\|_q$$

Now for the second part:

$$p = 1 \implies f^*(x) = \operatorname{sgn}(f(x)) \in L^\infty(E) \overset{f^* \text{ is sgn}}{\implies} \|f\|_\infty = 1 \implies f \in L^\infty(E)$$
$$\text{and } \int_E ff^* = \int_E |f| = \|f\|_1$$

Case 2. $p, q \neq 1$. Hölder's inequality is homogeneous (i.e. you can multiply or divide f and g by a constant and the inequality will still hold), therefore, without the loss of generality, we can assume that $\|f\|_p = \|g\|_q = 1$. Therefore, we need to prove that

$$\int_E |fg| \leqslant 1$$

Since $|f|^p$ and $|g|^q$ are integrable, it follows that f and g are infinite on subset of measure 0. Outside of that subset, by Young's inequality,

$$|fg| = |f| \cdot |g| \leqslant \frac{|f|^p}{p} + \frac{|g|^q}{q} \text{ almost everywhere on } E$$

Now let's put integrals on both sides:

$$\int_E |fg| \leqslant \int_E \frac{|f|^p}{p} + \int_E \frac{|g|^q}{q} = \left[\int_E |f|^p = \|f\|_p^p = 1^p = 1,\ \int_E |g|^q = \|g\|_q^q = 1^q = 1\right] = \frac{1}{p} + \frac{1}{q} = 1$$

Now let's do the second part.

$$ff^* \overset{\operatorname{sgn}(f) \text{ cancels out}}{=} \|f\|_p^{1-p} \cdot |f|^p \text{ on } E \implies \int_E ff^* = \|f\|_p^{1-p} \int_E |f|^p = \|f\|_p^{1-p} \cdot \|f\|_p^p = \|f\|_p$$
$$\int_E |f^*|^q = \|f\|_p^{(1-p)q} \int_E |f|^{(p-1)q} = [pq - p - q = 0] = \|f\|_p^{-p} \cdot \int_E |f|^p = \|f\|_p^{-p} \cdot \|f\|_p^p = 1$$
$$\int_E |f^*|^q = 1 \implies f^* \in L^q(E)$$

$\square$

Theorem 3 (Minkowski inequality)**.** E — measurable, $1 \leqslant p \leqslant \infty$. If $f, g \in L^p(E)$, then $f + g \in L^p$ and $\|f + g\|_p \leqslant \|f\|_p + \|g\|_p$

Proof.

Case 1. $p = 1$ — obvious, $p = \infty$ — exercise (use essential supremum).

Case 2. $p \in (1, +\infty)$. Assume that $[f+g] \neq [0]$. By part 2 of the previous theorem,

$$\|f+g\|_p = \int_E (f+g)(f+g)^* = \int_E f(f+g)^* + \int_E g(f+g)^* = (*)$$

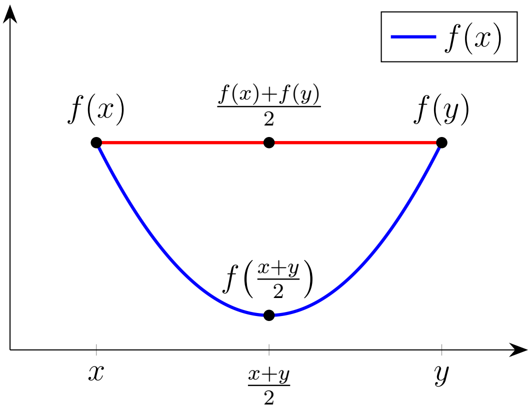
$f \in L^p(E),\ (f+g) \in L^p(E) \implies (f+g)^* \in L^q(E)$, now use Hölder's inequality:

$(*) = \|f\|_p \|(f+g)^*\|_q + \|g\|_p \|(f+g)^*\|_q = \|f\|_p + \|g\|_p$

Note that $f+g \in L^p$, as we proved earlier.

□

Remark. There's another way of looking at the Minkowski inequality — by considering convex functions. Let's use that to find a counterexample to the Minkowski inequality for $p < 1$.



A function f is convex if and only if for every x, y we have

$$f\left(\frac{x+y}{2}\right) \leqslant \frac{f(x)+f(y)}{2}$$
$$f(\text{average}) \leqslant \text{average}(f)$$

An integral is essentialy an average of an infinite number of points, therefore, we can write the same inequality in the integral form:

$$f\left(\int \varphi\right) \leqslant \int (f \circ \varphi)$$

If we define φ as

$$\varphi(x) = \begin{cases} a, & 0 \leqslant x \leqslant \frac{1}{2} \\ b, & \frac{1}{2} \leqslant x \leqslant 1 \end{cases}$$

Then we get the original inequality for two points.

Let $E = [0,1]$ and $f = c > 0,\ f = c = f_1 + f_2$. Then:

$$\|f\|_p = \left(\int_E |f|^p\right)^{\frac{1}{p}} = (c^p)^{\frac{1}{p}} = c$$

x^p is a concave function for $p < 1$. Therefore, the inequality will be in the other side:

$$\left(\int_E f_1\right)^p \geqslant \int_E f_1^p \implies \int_E f_1 \geqslant \|f_1\|_p$$
$$\left(\int_E f_2\right)^p \geqslant \int_E f_2^p \implies \int_E f_2 \geqslant \|f_2\|_p$$

If f_1 and f_2 are not constant, then the above inequalities are strict. Let's now sum up the two inequalities:

$$\begin{aligned}
&\|f_1\|_p + \|f_2\|_p < \int_E f_1 + \int_E f_2 = \int_E c = c = \|f\|_p \\
&\|f_1\|_p + \|f_2\|_p < \|f\|_p
\end{aligned}$$

Which is what we wanted.

Theorem 4 (Cauchy-Schwarz inequality)**.** If $f, g \in L^2$, then

$$\int fg \leqslant \sqrt{\int f^2 \cdot \int g^2}$$

In linear algebra, we prove that inequality for two vectors in $\mathbb{R}^n$ and their inner product. Therefore, we can define

$$\langle f, g\rangle \coloneqq \int fg$$

Then a lot of properties that hold for vectors will also hold for such an inner product.

Proof. This is Hölder's inequality for $p = q = 2$. □

Theorem 5. $m(E) < \infty$. If $1 \leqslant p_1 < p_2 \leqslant \infty$, then $L^{p_2}(E) \subset L^{p_1}(E)$.

Proof. Case $p_2 = \infty$ — exercise. Now assume that p_2 is finite. Let's define $p \coloneqq \frac{p_2}{p_1} > 1,\ q = \overline{p}$.
If $f \in L^{p_2}(E)$, then $|f|^{p_1} \in L^p(E)$. Let's define g as follows:

$$g \coloneqq \chi_E \implies \int_E g^q = \int_E 1 = m(E) < \infty \implies f \in L^q(E)$$

Here χ_E is the characteristic function of E. Now apply Hölder's inequality:

$$\int_E |f|^{p_1} = \int_E |f|^{p_1} g \leqslant \||f|^{p_1}\|_p \cdot \|g\|_q \overset{p_1 \cdot p = p_2}{=} \left(\int_E |f|^{p_2}\right) \cdot \left(\int_E |g|^q\right)^{\frac{1}{q}} \overset{g^q = g}{=} \|f\|_{p_2}^{p_1} \cdot \left(m(E)\right)^{\frac{1}{q}}$$

$\|f\|_{p_2}$ is finite, as $f \in L^{p_2}(E)$, and $m(E)$ is finite as well, therefore, the right side is finite, therefore, $f \in L^{p_1} E$. □

Remark.

Proposition 1. We have proved the non-strict inclusion. Let's prove that $L^{p_1}(E) \neq L^{p_2}(E)$, i.e. that the inclusion is actually strict.

Proof. Take $E = (0, 1]$, $f(x) = \frac{1}{x^\alpha}$. Let's take α such that $\alpha p_2 > 1$ and $\alpha p_1 < 1$ (which is possible as $p_1 < p_2$). Then:

$$\begin{aligned}
&\alpha p_2 > 1 \implies \int_E f^{p_2} = \int_0^1 \frac{1}{x^{\alpha p_2}} = \infty \implies f \notin L^{p_2} \\
&\alpha p_1 < 1 \implies \int_E f^{p_1} = \int_0^1 \frac{1}{x^{\alpha p_1}} < \infty \implies f \in L^{p_1} \\
&f \notin L^{p_2} \wedge f \in L^{p_1} \implies L^{p_2} \neq L^{p_1}
\end{aligned}$$

□

3.3 Separability of L^p

Theorem 1. E — measurable. If $1 \leqslant p < \infty$, then $L^p(E)$ is separable (i.e. there exists a countable dense subset in $L^p(E)$), and $L^\infty[a,b]$ is not separable.

Proposition 1. E — measurable, $1 \leqslant p \leqslant \infty$. Then the subspace of all simple functions is dense in $L^p(E)$.

Proposition 2. $[a,b]$ is a bounded interval, $1 \leqslant p < \infty$. Then the step functions are dense in the space of all simple functions on $[a,b]$. in $L^p([a,b])$.

After those propositions are proved, the proof of the theorem is going to go as follows. For $[a,b]$, let $S[a,b]$ be the set of all step functions with rational values and rational points of the subdivition. In this case, the set $S[a,b]$ is already countable. By those two propositions, $S[a,b]$ is dense in $L^p[a,b]$. Now, if E is arbitrary, we could consider the family

$$\mathcal{F} = \bigcup_{1}^{\infty} S[-n,n]$$

$\mathcal{F}$ is dense in $L^p(\mathbb{R})$. For any $E \subset \mathbb{R}$ restrict the functions from $\mathcal{F}$ to E.

Proof of Proposition 1.

1. $p = \infty$. Let $g \in L^\infty(E)$. We want to approximate g with a sequence of simple functions that converge to g. $g \in L^\infty(E)$, thus, g is bounded outside E_0, $m(E_0) = 0$. Now let's apply the simple approximation lemma for f restricted onto $E \setminus E_0$.

2. $p < \infty$. Then we can't use the simple approximation lemma directly, as we don't have essential boundness. Let's use the simple approximation theorem instead: there exists a sequence of simple functions $\{\varphi_n\}$, such that $\varphi_n \to g$ pointwise on E and $|\varphi_n(x)| \leqslant |g(x)|\ \forall n, x$. It follows that $\|\varphi_n\|_p \leqslant \|g\|_p < \infty \implies \varphi_n \in L^p(E)\ \forall n$.
   Since $|\varphi_n(x)| \leqslant |g(x)|$, we can also write that
   $$|\varphi_n - g|^p \leqslant (2|g|)^p = 2^p|g|^p$$
   Now we want to use the dominated convergence theorem. $g \in L^p(E)$, therefore, $|g|^p$ is integrable, and multiplying it by a constant (2^p) does not change that. So, we have dominated $|\varphi_n - g|^p$ with $2^p|g|^p$. Therefore, by dominated convergence theorem, we have:
   $$\lim_{n\to\infty} \int |\varphi_n - g|^p = \int \left| \lim_{n\to\infty} \varphi_n - g \right|^p = 0 \int |g-g|^p = 0 \implies \lim_{n\to\infty} \int |\varphi_n - g|^p = 0$$
   Therefore, $\varphi_n \to g$ in $L^p(E)$ by definition.

□

Proof of Proposition 2. A simple function can be represented as

$$\sum_{i=1}^{n} c_i \chi_{A_i}$$

Let's approximate each of the characteristic functions, and then take the linear combination with c_i as the coefficients.

So, we want to approximate χ_A by a step function. A is a measurable set of finite measure (since $[a, b]$ is bounded).

Therefore, as we have proved previously, for any $\varepsilon > 0$ there exists a finite disjoint collection of open intervals $\{I_k\}$, such that if $U = \cup I_k$, then $m(U \triangle A) < \varepsilon$. Let's choose $\varepsilon = \varepsilon^p$, then $m(U \triangle A) < \varepsilon^p$ instead. So:

$$\|\chi_A - \chi_U\|_p = \left(\int_{A \triangle U} 1^p\right) \leqslant (\varepsilon^p)^{\frac{1}{p}} \leqslant \varepsilon$$

□

Now let's prove that $L^\infty[a, b]$ is not separable.

Proof. Let $[a, b] = [0, 1]$.

For every $y \in (0, 1)$, let's define $f_y = \chi_{[0,y]}$. If $y_1 \neq y_2$, then $|f_{y_1}(x) - f_{y_2}(x)| = 1$ for every x between y_1 and y_2. Therefore, $\|f_{y_1} - f_{y_2}\|_\infty = 1$. If we now take balls in $L^\infty[a, b]$ of radius $\frac{1}{2}$ centered in those functions, they will not intersect (as the pairwise distances between centers are 1). Therefore, $L^\infty[a, b]$ is not separable, because we'd have to have at least one function in every one of those balls, but there's countably many of them. □

3.4 Completeness of L^p

***Definition* 1.** X — normed vector space.

Convergence: $f_n \to f$ if $\forall \varepsilon > 0 : \exists N : \forall n \geqslant N : \|f_n - f\| < \varepsilon$.

A sequence $\{f_n\}$ is *Cauchy* if $\forall \varepsilon > 0 : \exists N : \forall m, n \geqslant N : \|f_n - f_m\| < \varepsilon$

Proposition 1. X — normed space. Then:

1. Every convergent sequence is Cauchy.

2. If a Cauchy sequence has a convergent subsequence, then if converges.

Proof. We'll leave the proof as an exercise. The main idea is that once we can guess the limiting function f, we can usually prove the convergence by definition. □

***Definition* 2.** A sequence $\{f_n\}$ in X is called *rapidly Cauchy*, if there exists a convergent series of positive numbers $\sum_1^\infty \varepsilon_k$, such that

$$\|f_{k+1} - f_k\| \leqslant \varepsilon_k^2 \quad \forall k$$

Proposition 2. 1. Every rapidly Cauchy sequence is Cauchy.

2. Every Cauchy sequence has a rapidly Cauchy subsequence.

Proof. 1. If $\sum_1^\infty \varepsilon_k$ converges, then $\sum_1^\infty \varepsilon_k^2$ converges as well (as once the epsilons get less than one, that the squaring will only speed the convergence up).

Therefore,

$$\|f_n - f_m\| \leqslant \sum_n^{m-1} \varepsilon_k^2$$

and the tails of a convergent functions converge to 0.

2. Each next element should be chosen sufficiently far in the line, to ensure the consecutive differences are small enough.

□

Theorem 1. E — measurable, $1 \leqslant p \leqslant \infty$. Then every rapidly Cauchy sequence in $L^p(E)$ converges both with respect to $\|\cdot\|_p$ and pointwise almost everywhere on E to a function in $L^p(E)$.

Proof. The case $p = \infty$ will be in the homework. Assume that $p < \infty$, $\{f_n\}$ — rapidly Cauchy. Since our functions belong to the L^p space, every such function is finite except a subset of measure zero. Let's throw this subset of measure zero out for every one of our functions. Since there's a countable number of functions, the set that we will throw out will still have measure 0.

$\{f_n\}$ is rapidly Cauchy. Let's take the definition of a rapidly Cauchy sequence and exponentiate both sides to power p:

$$\|f_{k+1} - f_k\|_p = \left(\int_E |f_{k+1} - f_k|^p\right)^{\frac{1}{p}} \leqslant \varepsilon_k^2 \quad \forall k \implies \int_E |f_{k+1} - f_k|^p \leqslant \varepsilon_k^{2p} \quad \forall k$$

By Chebyshev,

$$m\{x \in E \mid |f_{k+1}(x) - f_k(x)| \geqslant \varepsilon_k\} = m\{x \in E \mid |f_{k+1}(x) - f_k(x)|^p \geqslant \varepsilon_k^p\} \leqslant \frac{1}{\varepsilon_k^p} \int_E |f_{k+1}(x) - f_k(x)|^p \leqslant \varepsilon_k^p$$

$p \geqslant 1$, $\sum_1^\infty \varepsilon_k < \infty \implies \sum_1^\infty \varepsilon_k^p < \infty$. Now we can use the Borel–Cantelli lemma. Let's say $E_0 \subset E$ is the subset of measure zero where Borel-Cantelli doesn't hold. Then every $x \in E \setminus E_0$ belongs to a finite number of such sets. Let's put $K(x)$ as the maximum index of such set plus one. Therefore:

$$|f_{k+1}(x) - f_k(x)| < \varepsilon_k \quad \forall k \geqslant K(x)$$

Therefore, for every $x \in E \setminus E_0$, $\{f_k(x)\}$ is a Cauchy sequence. As a result, f_k converges to some function f pointwise on $E \setminus E_0$.

Now we have a candidate for convergence, but we have to prove that it converges in the L_p space.

$$\int_E |f_{n+k} - f_n|^p \leqslant \sum_{m=n}^{\infty} \varepsilon_m^{2p} \quad \forall n, k$$

Take $k \to \infty$:

$$\int_E |f - f_n|^p \overset{\text{Fatou's lemma}}{\leqslant} \liminf_{k\to\infty} \int_E |f_{n+k} - f_n|^p$$
$$\|f - f_n\|_p \leqslant \left(\sum_{m=n}^{\infty} \varepsilon_m^{2p}\right)^{\frac{1}{p}}$$

This is the tail of a converging series, and thus it converges to 0. Therefore, $f_n \to f$ in the L^p norm.

Why is $f \in L^p$? That's true as L^p is a vector space, $f_n \in L^p, f - f_n \in L^p$. Then their sum is also in L^p: $f_n + (f - f_n) = f \in L^p$. □

Does pointwise convergence cause convergence in L^p, or vice versa? It turns out that neither is true.

Example 1.

$$f_n = n^{1/p} \cdot \chi_{(0,1/n]}, \;\; E = [0, 1]$$

$f_n \to 0$ pointwise. But:

$$\int_E |f_n - 0|^p = \int_0^{1/n} n = 1$$

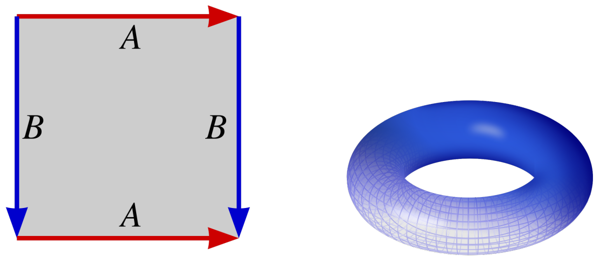
Example 2. Let's imagine a piano. When we're playing a piano, we're pressing the keys one by one from left to right. Let's split the interval of $[0, 1]$ into n pieces, and "move from left to right" by selecting functions equal to $\chi_{[k/n,\,(k+1)/n]}$. We can imagine that when we press a key (which is $[k/n,\ (k+1)/n]$) it is "lifted up" by 1 (as opposed to being pushed down as on a real piano).

When we've pressed every key from 1 to n from left to right, let's subdivide each key into two, so now we have $2n$ keys. Now let's press every key from 1 to $2n$ from left to right, and subdivide again, and so on.

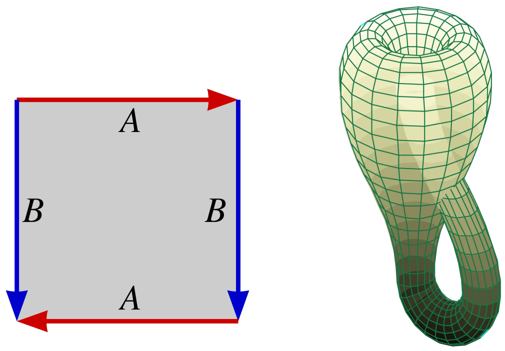
Each time the width of our piano keys will be smaller and smaller and smaller, and, therefore, the integral will converge to 0. However, there's no pointwise convergence on any subset of the interval, because every point on $[0, 1]$ will be "played" infinitely many times.

4 Manifolds

Let's first show some motivation for defining manifolds. Let's imagine that we have a rectangle cut out of paper that we're trying to fold. If we fold it and then glue the opposing sides as drawn on the figure, then this will be a torus:



However, it's not obvious whether we can glue the opposing sides like on the next figure. This is the so-called Klein bottle:



At every point of the surface it looks like $\mathbb{R}^2$!

4.1 Review of topology

***Definition* 1** (Topological space)**.** X — a set, $\tau = \{U_\alpha \subset X\}_{\alpha \in I}$ (I — some index set).

1. $\varnothing, X \in \tau$.
2. Arbitrary unions of sets from τ are in τ.
3. Finite intersections of sets from τ are in τ.

Then (X, τ) is a *topological space*. U_α are called *open sets*. Each U_α^C is a *closed set*.

Example. If $d(\cdot, \cdot)$ is a metric on X, then $B_r(x) = \{y \in X \mid d(x, y) < r\}$ and $U \subset X$ is open, if $\forall x \in U : \exists r > 0 : B_r(x) \subset U$.

A topological space is the smallest structure on a set that allows us to talk about convergence: a sequence of points in a set converges to a limit, if for every open set containing this limit the sequence is contained in that open set starting from some N.

***Definition* 2.** $f : X \to Y$ is *continuous* if the preimage of any open set in Y is open in X.

***Definition* 3.** A bijection $f : X \to Y$ with f and f^{-1} being continuous is called a *homeomorphism*.

Remark. Since f is continuous, this means that f^{-1} maps every open set in Y to an open set in X. But since f^{-1} is continuous as well, this means that f maps every open set in X to an open set in Y.

Therefore, f creates a bijection between open sets in X and in Y.

Example. 1. $\arctan : (-\infty, +\infty) \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ is a homeomorphism.

2.
$$f : [0,1) \cup [2,3] \to [0,2] \qquad f(x) = \begin{cases} x & \text{if } x \in [0,1) \\ x-1 & \text{if } x \in [2,3] \end{cases}$$

Here we have an induced topology on $[0,1) \cup [2,3]$. Is this a homeomorphism?

(a) f is a bijection.

(b) Is f continuous? We need to check that the preimage of an any open set U in $[0,2]$ is open in $[0,1) \cup [2,3]$. Part of the preimage will be mapped into $[0,1)$, and part will be mapped into $[2,3]$.
It can be shown that the left endpoint of $f^{-1}(U \cap [0,1))$ and the right endpoint of $f^{-1}(U \cap [1,2])$ will be excluded. Therefore, the preimage in question can be represented as

$$\left(f^{-1}(U \cap [0,1)),\ f^{-1}(U \cap [1,2])\right) \cap \left([0,1) \cup [2,3]\right)$$

As this is an intersection of an open set in $\mathbb{R}$ with our set, then it's open in the induced topology.

(c) f^{-1} is not continuous. Let's take the set $[2,3]$. It's open in $[0,1) \cup [2,3]$ by the definition of an induced topology. However,

$$(f^{-1})^{-1}([2,3]) = f([2,3]) = [1,2]$$

$[1,2]$ is not open, therefore, f^{-1} is not continuous.

***Definition* 4.** A topological space (X, τ) is called *Hausdorff* if for all $x, y \in X$, $x \neq y$ there are open neighborhoods U of x and V of y, such that $U \cap V = \varnothing$.

Example. Zariski topology on $\mathbb{R}$ or $\mathbb{C}$: U is open if and only if $U = \varnothing$ or $X \setminus U$ is finite.

It is not Hausdorff: there is no way to construct non-overlapping neighborhoods of two different points, because every open set is X without a just a finite number of points, which will still leave us with a continuum of common points.

***Definition* 5** (Basis)**.** X — set, $\mathcal{B}$ — collection of subsets of X, such that:

1.
$$X = \bigcup_{B \in \mathcal{B}} B, \quad \varnothing \in \mathcal{B}$$

2. For all B_1, $B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, there exists a $B_3 \in \mathcal{B}$ with $x \in B_3 \subset (B_1 \cap B_2)$.

Then the set of all unions of elements of $\mathcal{B}$ is called the topology generated by B.

Such $\mathcal{B}$ is called a *basis* of the topology it generates.

Proposition 1. The definition is correct, i.e. this is indeed a topology.

Proof. 1. $\varnothing$, $X \in \tau$ — true, because we can take the empty union and the union of all $B \in \mathcal{B}$.

2. Unions of U_α's from τ are in τ — we can just take the union of the corresponding elements from $\mathcal{B}$.

3. Let's prove that the intersection of two open sets U_α and U_β is open, then use induction to prove that for finite intersections.

   Let's assume $x \in U_\alpha \cap U_\beta$. Since U_α is a union of sets from $\mathcal{B}$, we can choose the one that contains x and call it B_1. Let's do the same thing with U_β and B_2. So, $x \in B_1 \cap B_2$ where $B_1 \subset U_1,\ B_2 \subset U_\beta$.

   By the definition of $\mathcal{B}$, there exists a $B_{3,x} \subset U_\alpha \cap U_\beta$, such that $x \in B_{3,x}$. $B_{3,x}$ is open, because it's a union of just one set from $\mathcal{B}$. Since $x \in U_\alpha \cap U_\beta$ was arbitrary,

$$U_\alpha \cap U_\beta = \bigcup_{x \in U_\alpha \cap U_\beta} B_{3,x}$$

   Therefore, $U_\alpha \cap U_\beta$ as a union of open sets.

□

Definition 6. A topological space is called *second-countable* if it has a countable basis.

Proposition 2. Second-countability implies separability.

Proof. Let's assume we have a second-countable set with a countable basis $\mathcal{B}$. Choose an arbitrary point in each element of $\mathcal{B}$ to form a countable set in A. If A is not dense, then there exists an open set U, such that $U \cap A = \varnothing$. But since $\mathcal{B}$ is a basis, U is some union of elements of $\mathcal{B}$. Let's take one of those elements, B_U. We've chosen a point x in B_U:

$$\begin{cases} x \in B_U \subset U \\ x \in A \end{cases} \implies x \in U \cap A \implies U \cap A \neq \varnothing ?!$$

□

Remark. The inverse is not true: there are separable sets that are not second-countable.

4.2 Manifolds

Definition 1. M is a *topological manifold of dimension* n if it is a Hausdorff and second-countable topological space, such that every point in M has a neighborhood homeomorphic to an open set in $\mathbb{R}^n$.

(For all $p \in M$ there exists an open $U \subset M$, such that $p \in U$ and there exist an open $V \in \mathbb{R}^n$ and a homeomorphism $\varphi : U \to V$.)

Remark. This definition would work for a Klein bottle. Locally around every point it looks like a plane.

Remark. It turns out that the number n is a topological invariant, i.e. two topological manifolds cannot be homeomorphic if their dimensions differ.

Exercise: reduce this to the following statement:

Theorem 1 (Brower's invariance of domain)**.** Let $U \subset \mathbb{R}^n$ be open, and let $f : U \to \mathbb{R}^n$ be continuous and injective. Then $f(U)$ is open in $\mathbb{R}^n$.

We're not gonna prove this theorem. Let's look at a simpler case: a continuous injective function in $\mathbb{R}$. $f : \mathbb{R} \to \mathbb{R}$. If $f \in C^1(\mathbb{R})$, then

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(|x - x_0|)$$

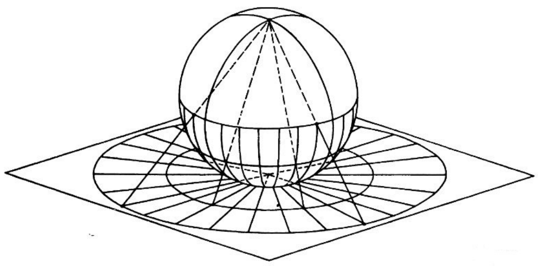
If $f'(x_0)$ is not zero, it follows that $f(x_0)$ cannot be a boundary point, since $f(x) - f(x_0)$ will have two different signs around x_0. But again, $f'(x_0)$ can have zeros, so this is not a full proof.

Definition 2. A pair (U, φ) with $U \subset M$ — an open set and $\varphi : U \to \mathbb{R}^n$ — a homeomorphism from U to $\varphi(U) = V$ is called a *chart*.

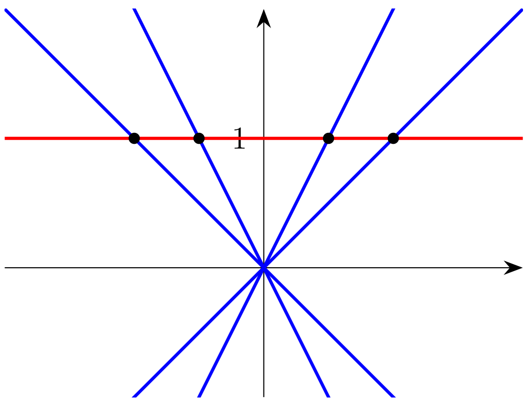
- φ is a local coordinate map.
- $\varphi(p) = \big(x^1(p), x^2(p), \dots, x^n(p)\big)$ — local coordinates. They don't work on the whole M — just on U.
- $\varphi^{-1} : V \to U$ is called a *coordinate system*.

Examples of topological manifolds:

1. An open set in $\mathbb{R}^n$. We only need one chart — $\mathbb{R}^n$ itself.
2. The n-sphere
$$S^n := \{(x^1, \dots, x^{n+1}) \in \mathbb{R}^{n+1} \mid \sum (x^j)^2 = 1\}$$
We would need two charts. Let's have one chart as the sphere except the north pole with homeomorphism as a stereographic projection, and the other as the sphere except the south pole. The two charts overlap, but they still satisfy the definition.

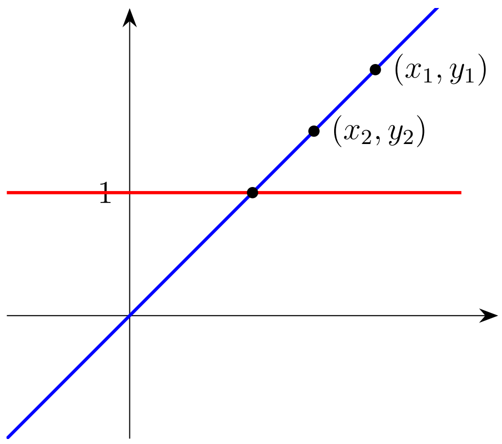


3. $\mathbb{P}^n$ — projective space. $\mathbb{P}^n = \big(\mathbb{R}^n \setminus \{0\}\big)/\sim$. (We're gonna define $\sim$ later).
$\mathbb{P}^1$ — the set of all lines in $\mathbb{R}^2$ passing through the origin:



Every non-horizontal line passing through the origin can be mapped onto a point on the red line. A horizontal line corresponds to the point at infinity. So, essentially, $\mathbb{P}^1$ is equivalent the extended real line.

Here's how we're gonna define our equivalence relation: $(x_1, y_1) \sim (x_2, y_2)$ if and only if $x_1 = kx_2$ and $y_1 = ky_2$, i.e. if they're on the same line:



This equivalence relation can be exteneded to $\mathbb{R}^{n+1}$.

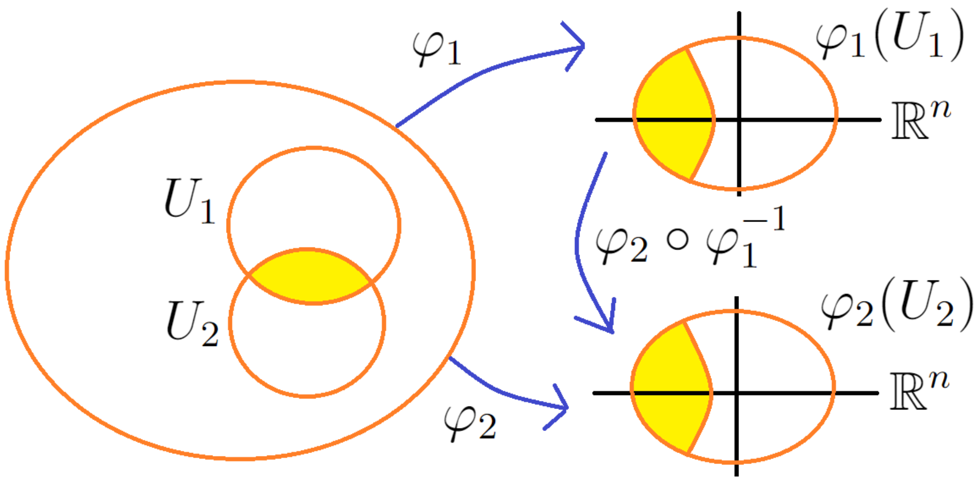
Now, here's how we're gonna define the charts. $p \in \mathbb{P}^n \implies$ take a representative $(x_1, \ldots, x_{n+1})$. Choose $x_i \neq 0$. Without the loss of generality, $i = n + 1$, therefore, in some neighborhood U around p we have $x_{n+1} \neq 0$. For each point in U:

$$\left(\frac{y_1}{y_{n+1}}, \frac{y_2}{y_{n+1}}, \ldots, \frac{y_n}{y_{n-1}}, 1\right) \sim (y_1, \ldots, y_{n+1})$$

So, around every point of the projective space we can find a chart: take any representative of that point, find a non-zero coordinate. This coordinate will remain non-zero in some neighborhood, therefore, we can divide by it, and this coordinate by which we divide will turn to 1. The remaining n coordinates will define the coordinates of a point in $\mathbb{R}^n$, which will provide a local homeomorphism around the point in the projective space to a subset $\mathbb{R}^n$, which defines a chart.

4.3 Smooth manifolds

Let's say we have two intersecting charts: U_1 and U_2 with maps φ_1 and φ_2. Each of those charts are mapped onto $\mathbb{R}^n$ by their respective local coordinate maps. If we consider $U_1 \cap U_2$, we can consider how it's mapped onto $\mathbb{R}^n$ by φ_1 and φ_2:



Since φ_1 and φ_2 are both homeomorphisms, $\varphi_2 \circ \varphi_1^{-1}$ is also a homeomorphism. It will convert the coordinates of $U_1 \cap U_2$ in the first and in the second chart.

***Definition* 1.** Let M be a topological manifold. Let $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}_{\alpha \in I}$, such that:

1. U_α are open and cover M.
2. For all α, β with $U_\alpha \cap U_\beta \neq \varnothing$, the *transition map*

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$$

   is C^r-smooth (i.e., it has r continuous derivatives).

Then $\mathcal{A}$ is called a C^r*-atlas* of M, and the pair $(M, \mathcal{A})$ is called a C^r*-manifold*.

Remark. This definition is different from the definition of a topological manifold, in that now we care about how we glue the charts together — we want it to be smooth. If we were to talk about smooth functions on a manifold, then we would run the computations separately on each chart. But in order for those computations to agree with each other, we would need the transition functions to be smooth as well.

Remark. By *smooth* we will mean C^∞-smooth.

***Definition* 2.** A C^r*-differentiable structure* on M is a maximal C^r-atlas on $\mathcal{A}$ (i.e., not contained in any other C^r-atlas).

***Definition* 3.** Two C^r-atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ are C^r*-equivalent*, if $\mathcal{A}_1 \cup \mathcal{A}_2$ is a C^r-atlas.

Proposition 1. C^r-equivalence is an equivalence relation.

Lemma. The union of all C^r-equivalent atlases in a single equivalence class is a maximal atlas. Hence, a maximal atlas exists.

Remark. We leave the proofs as an exercise.

There are atlases that don't admit a smooth structure! For $n = 1, 2, 3, 5, 6$, the n-dimensional sphere has only one smooth structure (i.e. C^∞-differentiable structure). If $n = 7$ — there are 28.

Examples of smooth manifolds:

1. $\mathbb{R}^n$ — it has a single chart.
2. Any open subset of a smooth manifold is a smooth manifold (we can intersect the charts with our subset).
3. A graph of a smooth function $f : \mathbb{R}^n \to \mathbb{R}$.
4. A sphere S^n. We can obtain at least one smooth structure using two stereographic projections (as before). The exercise is to check that the resulting conversion between those two charts is smooth.
5. A real projective space $\mathbb{P}^n$. We have defined $\mathbb{P}^n$ before as $(x_0, x_1, \ldots, x_n)/\sim$ with $n + 1$ different charts, depending on which coordinate of a given vector is non-zero.

Remark. The smooth structure does not define distances on the manifold.

4.4 Smooth maps between manifolds

TODO: insert picture.

***Definition* 1.** Let M, N be smooth manifolds. A map $f : M \to N$ is *smooth* at $p \in M$, if there exist charts (U, φ) with $p \in U$ and (V, ψ) with $f(p) \in V$, such that $f(U) \subset V$ and $\psi \circ f \circ \phi^{-1} : \varphi(U) \to \psi(V)$ is smooth (i.e. infinitely many times differentiable) at $\varphi(p)$.

***Remark*.** The key idea is that if we forget about the maps to $\mathbb{R}^m/\mathbb{R}^n$ and focus solely on M and N themselves, then we simply won't have enough structure to define the smoothness of a function.

***Definition* 2.** $f : M \to N$ is smooth, if it is smooth at every point of M.

***Remark*.** If $U \subset \mathbb{R}^m$, then $f : U \to \mathbb{R}^n$ is smooth if all partial derivatives exist.

Lemma. Smoothness of $f : M \to N$ is independent of the choice of charts in the definition, as long as $f(U) \subset V$.

Proof. We'll only give the proof idea. If we choose other charts, we'll get different ψ and ϕ in $\psi \circ f \circ \phi^{-1}$. In order to go back to the original ψ and ϕ, we'll have to compose with the transition maps, which are smooth because M and N are smooth manifolds, therefore the smoothness is preserved. □

Proposition 1. If M, N are smooth manifolds and $f : M \to N$ is smooth, then f is continuous (i.e. the preimage of an open set is open),

Proof. Let's prove it for any point $p \in M$. Let's suppose again that p is part of the chart (U, φ), and $f(p)$ is part of the chart (V, ψ). Then we can represent $f|_U$ as

$$f|_U = \varphi^{-1} \circ (\psi \circ f \circ \varphi^{-1}) \circ \varphi$$

φ^{-1} is continuous, as it is a homeomorphism. So is φ. $\psi \circ f \circ \varphi^{-1}$ is also continuous, because it is smooth in the sense of calculus (i.e. it's a map from $\mathbb{R}^m$ to $\mathbb{R}^n$). Therefore, $f|_U$ is continuous as a composition of three continuous functions.

Therefore, f is continuous at any point $p \in M$, which means that f is continuous. □

Proposition 2.

1. If $f : M \to N$ and $g : N \to P$ are smooth, then $g \circ f : M \to P$ is smooth.
2. If $f, g : M \to \mathbb{R}^n$, $\lambda : M \to \mathbb{R}$ are smooth, then $f + g$, λg and $\langle f, g\rangle = \sum f_i(g) \cdot g_i(x)$ are smooth functions.

Proof. This can be done by the direct application of the definition. □

***Definition* 3** (Diffeomorphism)**.** Let M, N be smooth manifolds. A bijection $f : M \to N$, such that f and f^{-1} are smooth is called a *diffeomorphism*.

Example. It can be proved that all compact connected manifolds of dimension 1 in $\mathbb{R}^2$ are diffeomorphic to a circle.

4.5 Tangent spaces

Model case

$M \subset \mathbb{R}^{n+1}$ is a surface given by the graph of a smooth map $F : \mathbb{R}^n \to \mathbb{R}$.

$$F(\vec{x} + \vec{v}) = F(\vec{x}) + \nabla F(\vec{x}) \cdot \vec{v} + \vec{o}(\|\vec{v}\|)$$

The first two terms are the linear part. They define the tangent space at the point $(\vec{x}, F(\vec{x}))$.

$$\nabla F = \left(\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_n}\right) \qquad \nabla F(\vec{x})\vec{v} = \frac{d}{dt}F(\vec{x}+t\vec{v})\Big|_{t=0} = D_{\vec{v}}F(\vec{x})$$

General case

Observe that for $f, g : \mathbb{R}^n \to \mathbb{R}$:

$$D_{\vec{v}}(fg)(x) = f(x)D_{\vec{v}}(g)(x) + g(x)D_{\vec{v}}(f)(x)$$

***Definition* 1.** Let M be a smooth manifold. Let $C^\infty(M)$ denote the space of all C^∞-smooth maps $f : M \to \mathbb{R}$.

Remark. $C^\infty(M)$ is a vector space (by Proposition 2), and it is not dependent on the choice of charts (by Lemma).

***Definition* 2** (Tangent space)**.** If $p \in M$, a linear map $D : C^\infty(M) \to \mathbb{R}$ is called a *derivation at p* if

$$D(fg) = f(p)D(g) + g(p)D(F)$$

for any $f, g \in C^\infty(M)$. Then $T_pM \coloneqq \{D : D \text{ is a derivation at } p\}$ is called the *tangent space* to M at p.

Lemma. Every derivation on $C^\infty(M \subset \mathbb{R}^n)$ at p is some directional derivative evaluated at p.