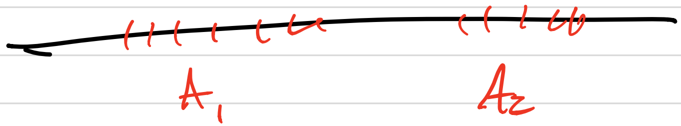
Recall that we introduced Lebesgue outer measure to generalise the idea of length and it does capture some of the properties of length.

There is one important property of length and volume that we have not looked at.

If A_1 and A_2 are disjoint subsets of $\mathbb{R}$ then we would like it to be the case that

$$\lambda^*(A_1 \cup A_2) = \lambda^*(A_1) + \lambda^*(A_2).$$

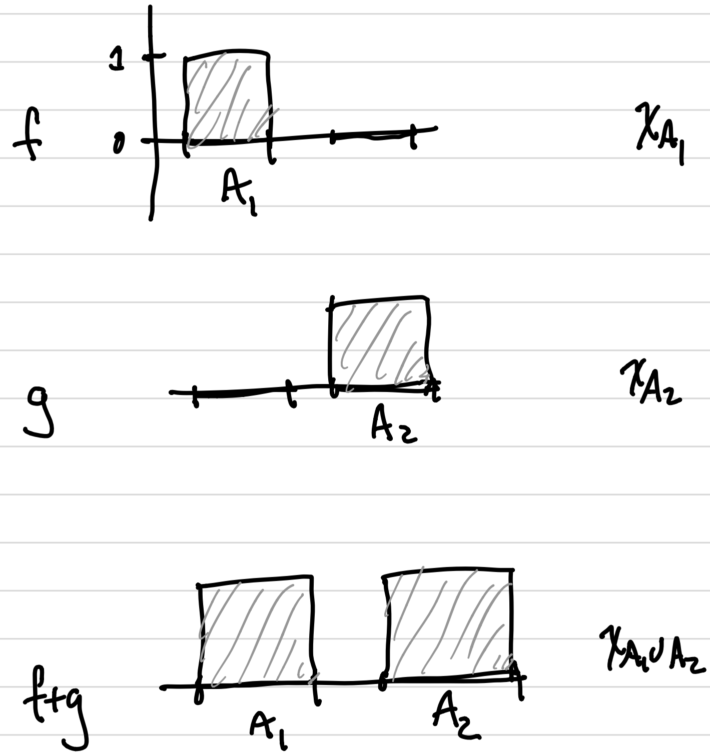


① When we construct the Lebesgue integral the additivity will be necessary in order to show that the integral is <u>linear</u>.

Suggestive picture:

Consider the "characteristic functions" of the sets.

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



$$\int f+g\,dx \overset{?}{=} \int f\,dx + \int g\,dx$$

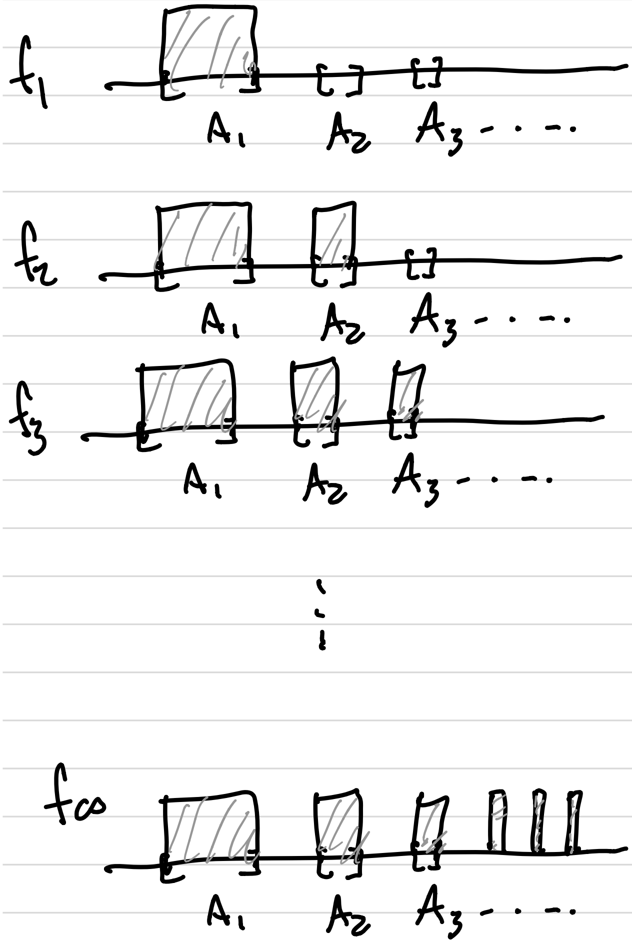
③ The condition we are most interested in is countable additivity (which is a strengthening of countable sub-additivity).

If $A_1, A_2 \ldots$ are pairwise disjoint then

$$\lambda^*\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \lambda^*(A_j).$$

This condition is required in order for the integral to have good limit properties:

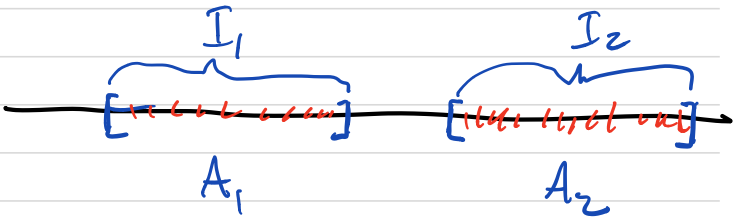
$$\int f_\infty \, dx \overset{?}{=} \lim_{n\to\infty} \int f_n \, dx$$

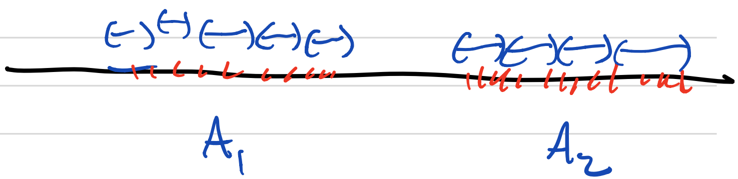


In order to get a picture of what is going on we will focus on a pair of sets.

② If A_1 and A_2 are contained in disjoint intervals then it will be the case that

$$\lambda^*(A_1 \cup A_2) = \lambda^*(A_1) + \lambda^*(A_2).$$





How would we prove that?

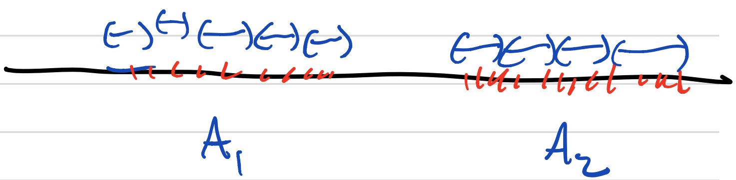
We know that if we take a cover of A_1 and a cover of A_2 and put them together we get a cover of $A_1 \cup A_2$. This shows that

$$\lambda^*(A_1 \cup A_2) \leq \lambda^*(A_1) + \lambda^*(A_2).$$

To get the other inequality:

$$\lambda^*(A_1 \cup A_2) \geq \lambda_1^*(A_1) + \lambda_1^*(A_2)$$

we start with an efficient cover of $A_1 \cup A_2$



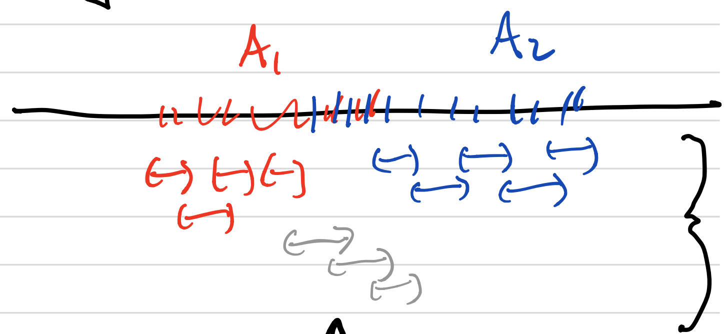
and show that we can use it to get covers of A_1 and A_2.

Note that in the picture each interval is either covering A_1 or A_2 so we can write $\mathcal{U}$ as $\mathcal{U}_1 \cup \mathcal{U}_2$ with

$$\underset{\lambda^*(A_1 \cup A_2)}{\underset{\text{SS}}{\sum_{u \in \mathcal{U}_1 \cup \mathcal{U}_2} \operatorname{length}(u)}} = \underset{\lambda^*(A_1)}{\underset{\text{VI}}{\sum_{u \in \mathcal{U}_1} \operatorname{length}(u)}} + \underset{\lambda^*(A_2)}{\underset{\text{VI}}{\sum_{u \in \mathcal{U}_2} \operatorname{length}(u)}}$$

How could additivity fail?

It could fail for the silly reason that A_1 and A_2 are not really disjoint:



Efficient cover of $A_1 \cup A_2$

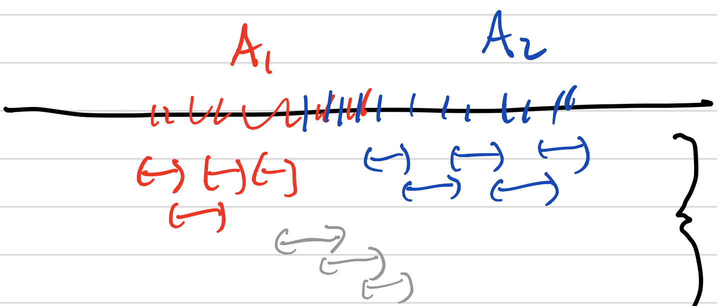
"indeterminant sets"

Here if we have an efficient cover of $A_1 \cup A_2$ we cannot reconstruct covers of A_1 and A_2 because we have intervals covering $A_1 \cap A_2$.

How could additivity fail? (part 2)

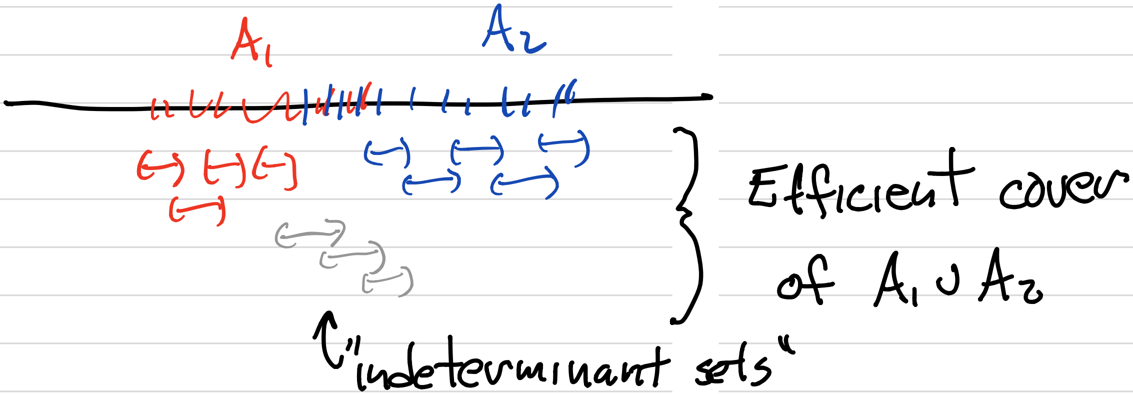
It could fail for the more subtle reason that even though A_1 and A_2 are disjoint

when we start with an efficient cover of $A_1 \cup A_2$ we have a substantial collection of intervals that cover both sets.



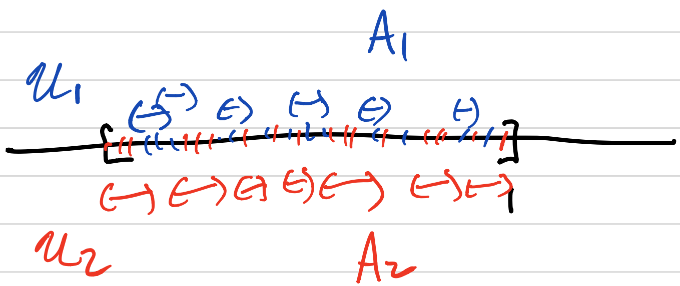
Efficient cover of $A_1 \cup A_2$

"indeterminant sets"



Here "substantial" means that the sum of the lengths of the "gray intervals" does not go to zero as our cover becomes more efficient.

To summarise it might be the case that A_1 and A_2 are so "intertwined" that every time we cover A_1 by $\mathcal{U}_1$ and cover A_2 by $\mathcal{U}_2$ the covers have substantial overlap.



So even though the sets are disjoint the covers are forced to overlap.

In particular this means that even though $\mathcal{U}_1 \cup \mathcal{U}_2$ is a cover of $A_1 \cup A_2$ it is not an efficient cover of $A_1 \cup A_2$.

In this discussion it turns out that for a given set $B \subset \mathbb{R}$ the set with which B is most "intertwined" is its complement.

We can define a set B to be measurable* if for any interval $I \subset \mathbb{R}$ it will be the case that

$$\lambda^*(B \cap I) + \lambda^*(B^c \cap I) = \text{length}(I).$$

Here we call the set I the "test set". Choosing different intervals lets us focus on different parts of B.

Intuitively this means that we have efficient covers of $B \cap I$ and $B^c \cap I$ which have small overlap so that when we put them together we get an efficient cover of I.

Why the asterisk?

This was Lebesgue's original definition of measurability.
It turns out that the particular collection of "test sets" that you choose does not play much of a role.

Caratheodory suggested a modification of Lebesgue's definition which makes it ① easier to work with and ② allows it to be applied to any outer measure and ③ is equivalent to Lebesgue's definition in the case of λ^*.

Def. Let μ^* be an outer measure defined on subsets of X.

We say $B \subset X$ is measurable if for every $A \subset X$ we

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

I

replacing length by outer measure

In other words we take our collection of test sets to be all subsets of X.

If

$$\mu^*(A \cap B) + \mu^*(A \cap B^c) = \mu^*(A),$$

we say that the pair $B \mid B^c$ give a "clean splitting" of A.

This means we can construct an efficient cover of A by putting together a cover of $A \cap B$ and a cover of $A \cap B^c$.

Let me give an example of a measurable set.

Proposition. If $\lambda^*(B)=0$ or $\lambda^*(B^c)=0$ then B is measurable.

Proof. Let's assume $\lambda^*(B)=0$. Let A be a "test set".

We want to show that

$$\lambda^*(A\cap B)+\lambda^*(A\cap B^c)=\lambda^*(A).$$

Sub-additivity of λ^* gives:

$$\lambda^*\big((A\cap B)\cup(A\cap B^c)\big)\leq \lambda^*(A\cap B)+\lambda^*(A\cap B^c)$$

$$\|$$

$$\lambda^*(A)$$

So we want to show the other inequality i.e.

$$\lambda^*(A \cap B) + \lambda^*(A \cap B^c) \leq \lambda^*(A).$$

$$\lambda^*(A\cap B)+\lambda^*(A\cap B^c)\le\lambda^*(B)+\lambda^*(A) = 0+\lambda^*(A) = \lambda^*(A).$$

since $\lambda^*(B)=0$

$A\cap B \subset B$, $A\cap B^c \subset A$

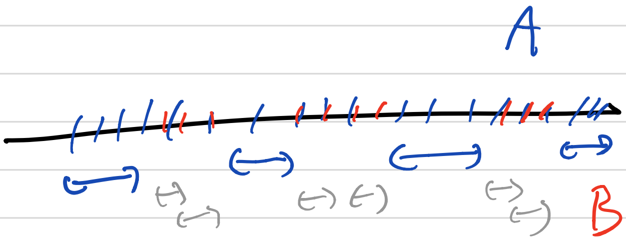
since $A\cap B\subset B$, $\lambda^*(A\cap B)\le\lambda^*(B)$

since $A\cap B^c\subset A$, $\lambda^*(A\cap B^c)\le\lambda^*(A)$.

$\lambda^*(A\cap B)\le\lambda^*(B)$ by sub-additivity.

$\lambda^*(B) = 0$ by assumption

$\lambda^*(A\cap B^c)\le\lambda^*(A)$ by sub-additivity.



If the gray intervals are "ambiguous" (covering A and B) there total length goes to zero as the cover gets more efficient.

Second example:

Proposition. The interval $(-\infty, b]$ is measurable.

Will prove this later.

It would be nice if all subsets of $\mathbb{R}$ were measurable for Lebesgue outer measure.

Unfortunately this is not the case.

So even though λ^* is defined for <u>all</u> subsets of $\mathbb{R}$ in order to achieve additivity we have to "throw some sets away."

(In fact we will have to restrict our consideration to measurable sets.)

This won't be so bad if this collection of sets contains all of the familiar sets and is closed under the standard operations that we use to build new sets from old.

Basic sets:

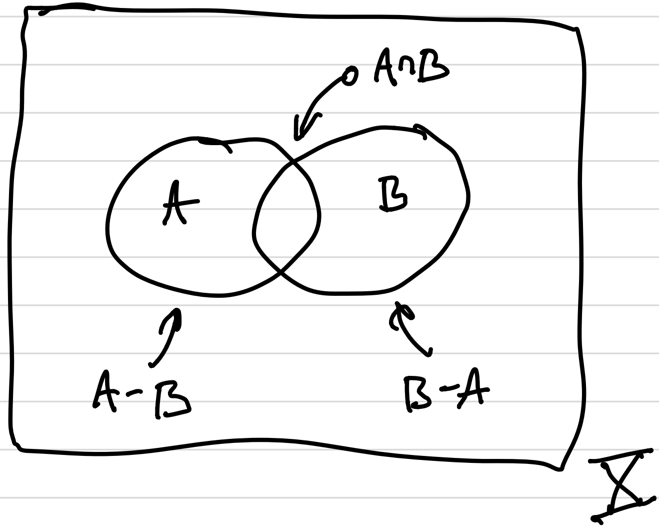
In $\mathbb{R}$ we might start with intervals: open, closed, half-open, half-infinite, finite sets, the integers, the rational numbers, irrational numbers...

Operations we want:

Countable unions

Countable intersections

Venn diagram operations



Venn diagram

$$A - B = A \cap B^c$$

$$B - A = B \cap A^c$$

The idea that a collection of sets is closed under standard operations is captured in the notion of a

σ- algebra.

(sigma)

A σ-algebra $\mathcal{A}$ is a collection of subsets of X such that

(a) $\phi \in \mathcal{A}$, $X \in \mathcal{A}$

(b) If $A \in \mathcal{A}$ then $A^c = X - A$ is in $\mathcal{A}$

(c) For each infinite sequence $\{A_i\}$ of sets in $\mathcal{A}$

$\bigcup_{i=1}^{\infty} A_i$ is in $\mathcal{A}$ and $\bigcap_{i=1}^{\infty} A_i$ is in $\mathcal{A}$.

An equivalent definition of a σ-algebra which is more efficient but less symmetric:

A σ-algebra $\mathcal{A}$ is a collection of subsets of X such that

(a) $X \in \mathcal{A}$ (since $\phi = X^c$)

(b) If $A \in \mathcal{A}$ then $A^c = X - A$ is in $\mathcal{A}$

(c) For each infinite sequence $\{A_i\}$ of sets in $\mathcal{A}$

$\bigcup_{i=1}^{\infty} A_i$ is in $\mathcal{A}$. (Since $\bigcap A_i = \left(\bigcup A_i^c\right)^c$)

The next idea we want to capture is what it means for a given collection of sets to generate a σ-algebra.

Think about vectors $v_1 \dots v_k$ in $\mathbb{R}^n$. There are two ways to describe the subspace V of $\mathbb{R}^n$ generated by $v_1 \dots v_k$.

① V is the set of all linear combinations $a_1 v_1 + \dots a_k v_k$.

② V is the smallest subspace of $\mathbb{R}^n$ that contains $v_1 \dots v_k$. That is to say that V is the intersection of all subspaces of $\mathbb{R}^n$ containing $v_1 \dots v_k$.

In our setting the first approach is too complicated.

We use the second approach.

Proposition. Given a collection of subsets of X, $\mathcal{F}$ there is a unique smallest σ-algebra containing $\mathcal{F}$.

Definition. If X is a <u>topological space</u> then the <u>Borel σ-algebra</u> is the σ-algebra generated by <u>open sets</u>.

We call a set in the Borel σ-algebra a <u>Borel set</u>.

Consider our favourite subsets of $\mathbb{R}$. I claim that they are in the σ-algebra generated by open sets.

Closed sets are Borel.

Points are Borel.

$\mathbb{Z}, \mathbb{N}$ are Borel.

$\mathbb{Q}$ is Borel.

$\mathbb{R}-\mathbb{Q}$ is Borel.

Theorem. Let μ^* be an outer measure on X. Then

1) the collection of μ^* measurable sets is a σ-algebra $\mathcal{M}$

2) the restriction of μ^* to $\mathcal{M}$ is countably additive.

Definition. Say X is a set, Σ is a σ-algebra of subsets of X and μ is a countably additive $[0,\infty]$ valued function on Σ then μ is a measure.

Definition. The restriction of λ^* to the measurable sets in $\mathbb{R}$ is called Lebesgue measure.

$$(\lambda^*, m, X)$$

Theorem. All Borel sets in $\mathbb{R}$ are measurable.

Thus <u>Lebesgue measure</u> is defined for all of our favourite sets.