

# Math212a1411

## Lebesgue measure.

Shlomo Sternberg

October 14, 2014

## Reminder

**No class this Thursday**

Today's lecture will be devoted to Lebesgue measure, a creation of Henri Lebesgue, in his thesis, one of the most famous theses in the history of mathematics.

# Henri Léon Lebesgue



**Born: 28 June 1875 in Beauvais, Oise, Picardie, France**

**Died: 26 July 1941 in Paris, France**

In today's lecture we will discuss the concept of *measurability* of a subset of  $\mathbb{R}$ . We will begin with Lebesgue's (1902) definition of measurability, which is easy to understand and intuitive. We will then give Caratheodory's (1914) definition of measurability which is highly non-intuitive but has great technical advantage. For subsets of  $\mathbb{R}$  these two definitions are equivalent (as we shall prove). But the Caratheodory definition extends to many much more general situations.

In particular, the Caratheodory definition will prove useful for us later, when we study Hausdorff measures.

# The $\epsilon/2^n$ trick.

An argument which will recur several times is:

$$\sum_1^{\infty} \frac{1}{2^n} = 1$$

and so

$$\sum_1^{\infty} \frac{\epsilon}{2^n} = \epsilon.$$

We will call this the “ $\epsilon/2^n$  trick” and not spell it out every time we use it.

- 1 Lebesgue outer measure.
- 2 Lebesgue inner measure.
- 3 Lebesgue's definition of measurability.
- 4 Caratheodory's definition of measurability.
- 5 Countable additivity.
- 6  $\sigma$ -fields, measures, and outer measures.
- 7 The Borel-Cantelli lemmas

# The definition of Lebesgue outer measure.

For **any** subset  $A \subset \mathbb{R}$  we define its **Lebesgue outer measure** by

$$m^*(A) := \inf \sum \ell(I_n) : I_n \text{ are intervals with } A \subset \bigcup I_n. \quad (1)$$

Here the length  $\ell(I)$  of any interval  $I = [a, b]$  is  $b - a$  with the same definition for half open intervals  $(a, b]$  or  $[a, b)$ , or open intervals.

Of course if  $a = -\infty$  and  $b$  is finite or  $+\infty$ , or if  $a$  is finite and  $b = +\infty$ , or if  $a = -\infty$  and  $b = \infty$  the length is infinite.



It doesn't matter if the intervals are open, half open or closed.

$$m^*(A) := \inf \sum \ell(I_n) : I_n \text{ are intervals with } A \subset \bigcup I_n. \quad (1)$$

The infimum in (1) is taken over all covers of  $A$  by intervals. By the  $\epsilon/2^n$  trick, i.e. by replacing each  $I_j = [a_j, b_j]$  by  $(a_j - \epsilon/2^{j+1}, b_j + \epsilon/2^{j+1})$  we may assume that the infimum is taken over open intervals. (Equally well, we could use half open intervals of the form  $[a, b)$ , for example.).

# Monotonicity, subadditivity.

If  $A \subset B$  then  $m^*(A) \leq m^*(B)$  since any cover of  $B$  by intervals is a cover of  $A$ .

For any two sets  $A$  and  $B$  we clearly have

$$m^*(A \cup B) \leq m^*(A) + m^*(B).$$

# Sets of measure zero don't matter.

A set  $Z$  is said to be of (Lebesgue) measure zero if its Lebesgue outer measure is zero, i.e. if it can be covered by a countable union of (open) intervals whose total length can be made as small as we like.

If  $Z$  is any set of measure zero, then  $m^*(A \cup Z) = m^*(A)$ .

# The outer measure of a finite interval is its length.

If  $A = [a, b]$  is an interval, then we can cover it by itself, so

$$m^*([a, b]) \leq b - a,$$

and hence the same is true for  $(a, b]$ ,  $[a, b)$ , or  $(a, b)$ . If the interval is infinite, it clearly can not be covered by a set of intervals whose total length is finite, since if we lined them up with end points touching they could not cover an infinite interval. We claim that

$$m^*(I) = \ell(I) \tag{2}$$

if  $I$  is a finite interval.

# Proof.

We may assume that  $I = [c, d]$  is a closed interval by what we have already said, and that the minimization in (1) is with respect to a cover by open intervals. So what we must show is that if

$$[c, d] \subset \bigcup_i (a_i, b_i)$$

then

$$d - c \leq \sum_i (b_i - a_i).$$

We first apply Heine-Borel to replace the countable cover by a finite cover. (This only decreases the right hand side of preceding inequality.) So let  $n$  be the number of elements in the cover.

We need to prove that if

$$[c, d] \subset \bigcup_{i=1}^n (a_i, b_i) \text{ then } d - c \leq \sum_{i=1}^n (b_i - a_i).$$

We do this by induction on  $n$ . If  $n = 1$  then  $a_1 < c$  and  $b_1 > d$  so clearly  $b_1 - a_1 > d - c$ .

Suppose that  $n \geq 2$  and we know the result for all covers (of all intervals  $[c, d]$ ) with at most  $n - 1$  intervals in the cover. If some interval  $(a_i, b_i)$  is disjoint from  $[c, d]$  we may eliminate it from the cover, and then we are in the case of  $n - 1$  intervals. So we may assume that every  $(a_i, b_i)$  has non-empty intersection with  $[c, d]$ .

Among the the intervals  $(a_i, b_i)$  there will be one for which  $a_i$  takes on the minimum possible value. By relabeling, we may assume that this is  $(a_1, b_1)$ . Since  $c$  is covered, we must have  $a_1 < c$ . If  $b_1 > d$  then  $(a_1, b_1)$  covers  $[c, d]$  and there is nothing further to do. So assume  $b_1 \leq d$ . We must have  $b_1 > c$  since  $(a_1, b_1) \cap [c, d] \neq \emptyset$ . Since  $b_1 \in [c, d]$ , at least one of the intervals  $(a_i, b_i)$ ,  $i > 1$  contains the point  $b_1$ . By relabeling, we may assume that it is  $(a_2, b_2)$ . But now we have a cover of  $[c, d]$  by  $n - 1$  intervals:

$$[c, d] \subset (a_1, b_2) \cup \bigcup_{i=3}^n (a_i, b_i).$$

So by induction  $d - c \leq (b_2 - a_1) + \sum_{i=3}^n (b_i - a_i)$ .

But  $b_2 - a_1 \leq (b_2 - a_2) + (b_1 - a_1)$  since  $a_2 < b_1$ .  $\square$

## We can use small intervals.

We repeat that the intervals used in (1) could be taken as open, closed or half open without changing the definition. If we take them all to be half open, of the form  $I_i = [a_i, b_i)$ , we can write each  $I_i$  as a disjoint union of finite or countably many intervals each of length  $< \epsilon$ . So it makes no difference to the definition if we also require the

$$\ell(I_i) < \epsilon \tag{3}$$

in (1). We will see that when we pass to other types of measures this will make a difference.



# Summary of where we are so far.

We have verified, or can easily verify the following properties:

1

$$m^*(\emptyset) = 0.$$

2  $A \subset B \Rightarrow m^*(A) \leq m^*(B).$

3 By the  $\epsilon/2^n$  trick, for any finite or countable union we have

$$m^*\left(\bigcup_i A_i\right) \leq \sum_i m^*(A_i).$$

4 If  $\text{dist}(A, B) > 0$  then

$$m^*(A \cup B) = m^*(A) + m^*(B).$$

5  $m^*(A) = \inf\{m^*(U) : U \supset A, U \text{ open}\}.$

6 For an interval

$$m^*(I) = \ell(I).$$

The only items that we have not done already are items 4 and 5. But these are immediate: for 4 we may choose the intervals in (1) all to have length  $< \epsilon$  where  $2\epsilon < \text{dist}(A, B)$  so that there is no overlap. As for item 5, we know from 2 that  $m^*(A) \leq m^*(U)$  for any set  $U \supset A$ , in particular for any open set  $U$  which contains  $A$ . We must prove the reverse inequality: if  $m^*(A) = \infty$  this is trivial. Otherwise, we may take the intervals in (1) to be open and then the union on the right is an open set whose Lebesgue outer measure is less than  $m^*(A) + \delta$  for any  $\delta > 0$  if we choose a close enough approximation to the infimum.

# Lebesgue outer measure on $\mathbb{R}^n$ .

All the above works for  $\mathbb{R}^n$  instead of  $\mathbb{R}$  if we replace the word “interval” by “rectangle”, meaning a rectangular parallelepiped, i.e a set which is a product of one dimensional intervals. We also replace length by volume (or area in two dimensions). What is needed is the following:

## Lemma

*Let  $\mathcal{C}$  be a finite non-overlapping collection of closed rectangles all contained in the closed rectangle  $J$ . Then*

$$\text{vol } J \geq \sum_{I \in \mathcal{C}} \text{vol } I.$$

*If  $\mathcal{C}$  is any finite collection of rectangles such that*

$$J \subset \bigcup_{I \in \mathcal{C}} I$$

*then*

$$\text{vol } J \leq \sum_{I \in \mathcal{C}} \text{vol } (I).$$

This lemma occurs on page 1 of Stroock, *A concise introduction to the theory of integration* together with its proof. I will take this for granted. In the next few slides I will talk as if we are in  $\mathbb{R}$ , but everything goes through unchanged if  $\mathbb{R}$  is replaced by  $\mathbb{R}^n$ .

In fact, once we will have developed enough abstract theory, we will not need this lemma. So for those of you who are purists and do not want to go through the lemma in Stroock, just stick to  $\mathbb{R}$ .

# The definition of Lebesgue inner measure.

Item 5. in our list said that the Lebesgue outer measure of any set is obtained by approximating it from the outside by open sets. The **Lebesgue inner measure** is defined as

$$m_*(A) = \sup\{m^*(K) : K \subset A, K \text{ compact}\}. \quad (4)$$

Clearly

$$m_*(A) \leq m^*(A)$$

since  $m^*(K) \leq m^*(A)$  for any  $K \subset A$ . We also have

# The Lebesgue inner measure of an interval.

Proposition.

For any interval  $I$  we have

$$m_*(I) = \ell(I). \quad (5)$$

## Proof.

If  $\ell(I) = \infty$  the result is obvious. So we may assume that  $I$  is a finite interval which we may assume to be open,  $I = (a, b)$ . If  $K \subset I$  is compact, then  $I$  is a cover of  $K$  and hence from the definition of outer measure  $m^*(K) \leq \ell(I)$ . So  $m_*(I) \leq \ell(I)$ . On the other hand, for any  $\epsilon > 0$ ,  $\epsilon < \frac{1}{2}(b - a)$  the interval  $[a + \epsilon, b - \epsilon]$  is compact and  $m^*([a - \epsilon, a + \epsilon]) = b - a - 2\epsilon \leq m_*(I)$ . Letting  $\epsilon \rightarrow 0$  proves the proposition.  $\square$



## Lebesgue's definition of measurability for sets of finite outer measure.

A set  $A$  with  $m^*(A) < \infty$  is said to be **measurable in the sense of Lebesgue** if

$$m_*(A) = m^*(A). \quad (6)$$

If  $A$  is measurable in the sense of Lebesgue, we write

$$m(A) = m_*(A) = m^*(A). \quad (7)$$

If  $K$  is a compact set, then  $m_*(K) = m^*(K)$  since  $K$  is a compact set contained in itself. Hence all compact sets are measurable in the sense of Lebesgue. If  $I$  is a bounded interval, then  $I$  is measurable in the sense of Lebesgue by the preceding Proposition.

# Lebesgue's definition of measurability for sets of infinite outer measure.

If  $m^*(A) = \infty$ , we say that  $A$  is measurable in the sense of Lebesgue if all of the sets  $A \cap [-n, n]$  are measurable.

# Countable additivity.

## Theorem

*If  $A = \bigcup A_i$  is a (finite or) countable disjoint union of sets which are measurable in the sense of Lebesgue, then  $A$  is measurable in the sense of Lebesgue and*

$$m(A) = \sum_i m(A_i).$$

In the proof we may assume that  $m(A) < \infty$  - otherwise apply the result to  $A \cap [-n, n]$  and  $A_i \cap [-n, n]$  for each  $n$ .

We have

$$m^*(A) \leq \sum_n m^*(A_n) = \sum_n m(A_n).$$

Let  $\epsilon > 0$ , and for each  $n$  choose compact  $K_n \subset A_n$  with

$$m^*(K_n) \geq m_*(A_n) - \frac{\epsilon}{2^n} = m(A_n) - \frac{\epsilon}{2^n}$$

which we can do since  $A_n$  is measurable in the sense of Lebesgue. The sets  $K_n$  are pairwise disjoint, hence, being compact, at positive distances from one another. Hence

$$m^*(K_1 \cup \cdots \cup K_n) = m^*(K_1) + \cdots + m^*(K_n)$$

and  $K_1 \cup \cdots \cup K_n$  is compact and contained in  $A$ .

$$m^*(K_n) \geq m_*(A_n) - \frac{\epsilon}{2^n} = m(A_n) - \frac{\epsilon}{2^n}$$

$$m^*(K_1 \cup \cdots \cup K_n) = m^*(K_1) + \cdots + m^*(K_n)$$

and  $K_1 \cup \cdots \cup K_n$  is compact and contained in  $A$ . Hence

$$m_*(A) \geq m^*(K_1) + \cdots + m^*(K_n),$$

and since this is true for all  $n$  we have

$$m_*(A) \geq \sum_n m(A_n) - \epsilon.$$

Since this is true for all  $\epsilon > 0$  we get

$$m_*(A) \geq \sum m(A_n).$$

But then  $m_*(A) \geq m^*(A)$  and so they are equal, so  $A$  is measurable in the sense of Lebesgue, and  $m(A) = \sum m(A_i)$ . □

# Open sets are measurable in the sense of Lebesgue.

## Proof.

Any open set  $O$  can be written as the countable union of open intervals  $I_i$ , and

$$J_n := I_n \setminus \bigcup_{i=1}^{n-1} I_i$$

is a disjoint collection of intervals (some open, some closed, some half open) and  $O$  is the disjoint union of the  $J_n$ . So every open set is a disjoint union of intervals hence measurable in the sense of Lebesgue. □

# Closed sets are measurable in the sense of Lebesgue.

This requires a bit more work:

If  $F$  is closed, and  $m^*(F) = \infty$ , then  $F \cap [-n, n]$  is compact, and so  $F$  is measurable in the sense of Lebesgue. Suppose that

$$m^*(F) < \infty.$$

For any  $\epsilon > 0$  consider the sets

$$G_{1,\epsilon} := \left[-1 + \frac{\epsilon}{2^2}, 1 - \frac{\epsilon}{2^2}\right] \cap F$$

$$G_{2,\epsilon} := \left([-2 + \frac{\epsilon}{2^3}, -1] \cap F\right) \cup \left([1, 2 - \frac{\epsilon}{2^3}] \cap F\right)$$

$$G_{3,\epsilon} := \left([-3 + \frac{\epsilon}{2^4}, -2] \cap F\right) \cup \left([2, 3 - \frac{\epsilon}{2^4}] \cap F\right)$$

$$\vdots$$

and set

$$G_\epsilon := \bigcup_i G_{i,\epsilon}.$$

The  $G_{i,\epsilon}$  are all compact, and hence measurable in the sense of Lebesgue, and the union in the definition of  $G_\epsilon$  is disjoint, so  $G_\epsilon$  is measurable in the sense of Lebesgue.



The  $G_{i,\epsilon}$  are all compact, and hence measurable in the sense of Lebesgue, and the union in the definition of  $G_\epsilon$  is disjoint, so is measurable in the sense of Lebesgue. Furthermore, the sum of the lengths of the “gaps” between the intervals that went into the definition of the  $G_{i,\epsilon}$  is  $\epsilon$ . So

$$m(G_\epsilon) + \epsilon = m^*(G_\epsilon) + \epsilon \geq m^*(F) \geq m^*(G_\epsilon) = m(G_\epsilon) = \sum_i m(G_{i,\epsilon}).$$

In particular, the sum on the right converges, and hence by considering a finite number of terms, we will have a finite sum whose value is at least  $m(G_\epsilon) - \epsilon$ . The corresponding union of sets will be a compact set  $K_\epsilon$  contained in  $F$  with

$$m(K_\epsilon) \geq m^*(F) - 2\epsilon.$$

Hence all closed sets are measurable in the sense of Lebesgue. □

# The inner-outer characterization of measurability.

## Theorem

*$A$  is measurable in the sense of Lebesgue if and only if for every  $\epsilon > 0$  there is an open set  $U \supset A$  and a closed set  $F \subset A$  such that*

$$m(U \setminus F) < \epsilon.$$

**Proof in one direction.** Suppose that  $A$  is measurable in the sense of Lebesgue with  $m(A) < \infty$ . Then there is an open set  $U \supset A$  with  $m(U) < m^*(A) + \epsilon/2 = m(A) + \epsilon/2$ , and there is a compact set  $F \subset A$  with  $m(F) \geq m_*(A) - \epsilon/2 = m(A) - \epsilon/2$ . Since  $U \setminus F$  is open, it is measurable in the sense of Lebesgue, and so is  $F$  as it is compact. Also  $F$  and  $U \setminus F$  are disjoint. Hence

$$m(U \setminus F) = m(U) - m(F) < m(A) + \frac{\epsilon}{2} - \left(m(A) - \frac{\epsilon}{2}\right) = \epsilon.$$

If  $A$  is measurable in the sense of Lebesgue, and  $m(A) = \infty$ , we can apply the above to  $A \cap I$  where  $I$  is any compact interval, in particular to the interval  $[-n, n]$ . So there exist open sets  $U_n \supset A \cap [-n, n]$  and closed sets  $F_n \subset A \cap [-n, n]$  with  $m(U_n/F_n) < \epsilon/2^n$ . Let  $U := \bigcup U_n$  and

$$F := \bigcup (F_n \cap ([-n, -n+1] \cup [n-1, n])).$$

A convergent sequence in  $F$  must eventually lie in an interval of length one and hence to at most the union of two entries in the union defining  $F$ , so  $F$  is closed. We have  $U \supset A \supset F$  and

$$U/F \subset \bigcup (U_n/F_n).$$

Indeed, any  $p \in U$  must belong to some  $([-n, -n+1] \cup [n-1, n])$ , and so if it does not belong to  $F$  then it does not belong to  $F_n$ . So

$$m(U/F) \leq \sum m(U_n/F_n) < \epsilon. \quad \square$$

## Proof in the other direction.

Suppose that for each  $\epsilon$ , there exist  $U \supset A \supset F$  with  $m(U \setminus F) < \epsilon$ . Suppose that  $m^*(A) < \infty$ . Then  $m(F) < \infty$  and  $m(U) \leq m(U \setminus F) + m(F) < \epsilon + m(F) < \infty$ . Then

$$m^*(A) \leq m(U) < m(F) + \epsilon = m_*(F) + \epsilon \leq m_*(A) + \epsilon.$$

Since this is true for every  $\epsilon > 0$  we conclude that  $m_*(A) \geq m^*(A)$  so they are equal and  $A$  is measurable in the sense of Lebesgue.

If  $m^*(A) = \infty$ , we have

$U \cap (-n - \epsilon, n + \epsilon) \supset A \cap [-n, n] \supset F \cap [-n, n]$  and

$$m((U \cap (-n - \epsilon, n + \epsilon)) \setminus (F \cap [-n, n])) < 2\epsilon + \epsilon = 3\epsilon$$

so we can proceed as before to conclude that

$$m_*(A \cap [-n, n]) = m^*(A \cap [-n, n]). \quad \square$$

# Consequences of the theorem.

## Proposition.

If  $A$  is measurable in the sense of Lebesgue, so is its complement  $A^c = \mathbb{R} \setminus A$ .

## Proof.

Indeed, if  $F \subset A \subset U$  with  $F$  closed and  $U$  open, then  $F^c \supset A^c \supset U^c$  with  $F^c$  open and  $U^c$  closed. Furthermore,  $F^c \setminus U^c = U \setminus F$  so if  $A$  satisfies the condition of the theorem so does  $A^c$ . □

## More consequences.

### Proposition.

If  $A$  and  $B$  are measurable in the sense of Lebesgue so is  $A \cap B$

### Proof.

For  $\epsilon > 0$  choose  $U_A \supset A \supset F_A$  and  $U_B \supset B \supset F_B$  with  $m(U_A \setminus F_A) < \epsilon/2$  and  $m(U_B \setminus F_B) < \epsilon/2$ . Then

$$(U_A \cap U_B) \supset (A \cap B) \supset (F_A \cap F_B)$$

and

$$(U_A \cap U_B) \setminus (F_A \cap F_B) \subset (U_A \setminus F_A) \cup (U_B \setminus F_B).$$



## Still more consequences of the theorem.

### Proposition.

If  $A$  and  $B$  are measurable in the sense of Lebesgue then so is  $A \cup B$ .

### Proof.

Indeed,  $A \cup B = (A^c \cap B^c)^c$ . □

Since  $A \setminus B = A \cap B^c$  we also get

### Proposition.

If  $A$  and  $B$  are measurable in the sense of Lebesgue then so is  $A \setminus B$ .

## Caratheodory's definition of measurability.

A set  $E \subset \mathbb{R}$  is said to be **measurable according to Caratheodory** if for any set  $A \subset \mathbb{R}$  we have

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c) \quad (8)$$

where (recall that)  $E^c$  denotes the complement of  $E$ . In other words,  $A \cap E^c = A \setminus E$ . This definition has many advantages, as we shall see. Our first task will be to show that it is equivalent to Lebesgue's.

Notice that we always have  $m^*(A) \leq m^*(A \cap E) + m^*(A \setminus E)$  so condition (8) is equivalent to

$$m^*(A \cap E) + m^*(A \setminus E) \leq m^*(A) \quad (9)$$

for all  $A$ .



# Caratheodory = Lebesgue.

## Theorem

*A set  $E$  is measurable in the sense of Caratheodory if and only if it is measurable in the sense of Lebesgue.*

**Proof that Lebesgue  $\Rightarrow$  Caratheodory.** Suppose  $E$  is measurable in the sense of Lebesgue. Let  $\epsilon > 0$ . Choose  $U \supset E \supset F$  with  $U$  open,  $F$  closed and  $m(U \setminus F) < \epsilon$  which we can do by the “inner-outer” Theorem. Let  $V$  be an open set containing  $A$ . Then  $A \setminus E \subset V \setminus F$  and  $A \cap E \subset (V \cap U)$  so

$A \setminus E \subset V \setminus F$  and  $A \cap E \subset (V \cap U)$  so

$$\begin{aligned} m^*(A \setminus E) + m^*(A \cap E) &\leq m(V \setminus F) + m(V \cap U) \\ &\leq m(V \setminus U) + m(U \setminus F) + m(V \cap U) \\ &\leq m(V) + \epsilon. \end{aligned}$$

(We can pass from the second line to the third since both  $V \setminus U$  and  $V \cap U$  are measurable in the sense of Lebesgue and we can apply the proposition about disjoint unions.) Taking the infimum over all open  $V$  containing  $A$ , the last term becomes  $m^*(A) + \epsilon$ , and as  $\epsilon$  is arbitrary, we have established (9) showing that  $E$  is measurable in the sense of Caratheodory.

## Caratheodory $\Rightarrow$ Lebesgue, case where $m^*(E) < \infty$ .

Then for any  $\epsilon > 0$  there exists an open set  $U \supset E$  with  $m(U) < m^*(E) + \epsilon$ . We may apply condition (8) to  $A = U$  to get

$$m(U) = m^*(U \cap E) + m^*(U \setminus E) = m^*(E) + m^*(U \setminus E)$$

so

$$m^*(U \setminus E) < \epsilon.$$

This means that there is an open set  $V \supset (U \setminus E)$  with  $m(V) < \epsilon$ . But we know that  $U \setminus V$  is measurable in the sense of Lebesgue, since  $U$  and  $V$  are, and

$$m(U) \leq m(V) + m(U \setminus V)$$

so

$$m(U \setminus V) > m(U) - \epsilon.$$

$$m(U \setminus V) > m(U) - \epsilon.$$

So there is a closed set  $F \subset U \setminus V$  with  $m(F) > m(U) - \epsilon$ . But since  $V \supset U \setminus E = U \cap E^c$ , we have

$$U \setminus V = U \cap V^c \subset U \cap (U^c \cup E) = E.$$

So  $F \subset E$ . So  $F \subset E \subset U$  and

$$m(U \setminus F) = m(U) - m(F) < \epsilon.$$

Hence  $E$  is measurable in the sense of Lebesgue.

# Caratheodory $\Rightarrow$ Lebesgue, case where $m^*(E) = \infty$ .

If  $m(E) = \infty$ , we must show that  $E \cap [-n, n]$  is measurable in the sense of Caratheodory, for then it is measurable in the sense of Lebesgue from what we already know. We know that the interval  $[-n, n]$  itself, being measurable in the sense of Lebesgue, is measurable in the sense of Caratheodory. So we will have completed the proof of the theorem if we show that the intersection of  $E$  with  $[-n, n]$  is measurable in the sense of Caratheodory.

# Unions and intersections.

More generally, we will show that the union or intersection of two sets which are measurable in the sense of Caratheodory is again measurable in the sense of Caratheodory. Notice that the definition (8) is symmetric in  $E$  and  $E^c$  so if  $E$  is measurable in the sense of Caratheodory so is  $E^c$ . So it suffices to prove the next lemma to complete the proof.

## Lemma

*If  $E_1$  and  $E_2$  are measurable in the sense of Caratheodory so is  $E_1 \cup E_2$ .*

# Proof of the lemma, 1.

For any set  $A$  we have

$$m^*(A) = m^*(A \cap E_1) + m^*(A \cap E_1^c)$$

by (8) applied to  $E_1$ . Applying (8) to  $A \cap E_1^c$  and  $E_2$  gives

$$m^*(A \cap E_1^c) = m^*(A \cap E_1^c \cap E_2) + m^*(A \cap E_1^c \cap E_2^c).$$

Substituting this back into the preceding equation gives

$$m^*(A) = m^*(A \cap E_1) + m^*(A \cap E_1^c \cap E_2) + m^*(A \cap E_1^c \cap E_2^c). \quad (10)$$

Since  $E_1^c \cap E_2^c = (E_1 \cup E_2)^c$  we can write this as

$$m^*(A) = m^*(A \cap E_1) + m^*(A \cap E_1^c \cap E_2) + m^*(A \cap (E_1 \cup E_2)^c).$$

## Proof of the lemma, 2.

$$m^*(A) = m^*(A \cap E_1) + m^*(A \cap E_1^c \cap E_2) + m^*(A \cap (E_1 \cup E_2)^c).$$

Now  $A \cap (E_1 \cup E_2) = (A \cap E_1) \cup (A \cap (E_1^c \cap E_2))$  so

$$m^*(A \cap E_1) + m^*(A \cap E_1^c \cap E_2) \geq m^*(A \cap (E_1 \cup E_2)).$$

Substituting this for the two terms on the right of the previous displayed equation gives

$$m^*(A) \geq m^*(A \cap (E_1 \cup E_2)) + m^*(A \cap (E_1 \cup E_2)^c)$$

which is just (9) for the set  $E_1 \cup E_2$ . This proves the lemma and the theorem.  $\square$



# The class $\mathcal{M}$ of measurable sets.

We let  $\mathcal{M}$  denote the class of measurable subsets of  $\mathbb{R}$  - “measurability” in the sense of Lebesgue or Caratheodory these being equivalent. Notice by induction starting with two terms as in the lemma, that any finite union of sets in  $\mathcal{M}$  is again in  $\mathcal{M}$

The first main theorem in the subject is the following description of  $\mathcal{M}$  and the function  $m$  on it:

### Theorem

*$\mathcal{M}$  and the function  $m : \mathcal{M} \rightarrow [0, \infty]$  have the following properties:*

- $\mathbb{R} \in \mathcal{M}$ .
- $E \in \mathcal{M} \Rightarrow E^c \in \mathcal{M}$ .
- If  $E_n \in \mathcal{M}$  for  $n = 1, 2, 3, \dots$  then  $\bigcup_n E_n \in \mathcal{M}$ .
- If  $F_n \in \mathcal{M}$  and the  $F_n$  are pairwise disjoint, then  $F := \bigcup_n F_n \in \mathcal{M}$  and

$$m(F) = \sum_{n=1}^{\infty} m(F_n).$$

We already know the first two items on the list, and we know that a finite union of sets in  $\mathcal{M}$  is again in  $\mathcal{M}$ . We also know the last assertion. But it will be instructive and useful for us to have a proof starting directly from Caratheodory's definition of measurability:

Recall

$$m^*(A) = m^*(A \cap E_1) + m^*(A \cap E_1^c \cap E_2) + m^*(A \cap E_1^c \cap E_2^c). \quad (10)$$

If  $F_1 \in \mathcal{M}$ ,  $F_2 \in \mathcal{M}$  and  $F_1 \cap F_2 = \emptyset$  then taking

$$A = F_1 \cup F_2, \quad E_1 = F_1, \quad E_2 = F_2$$

in (10) gives

$$m(F_1 \cup F_2) = m(F_1) + m(F_2).$$

Induction then shows that if  $F_1, \dots, F_n$  are pairwise disjoint elements of  $\mathcal{M}$  then their union belongs to  $\mathcal{M}$  and

$$m(F_1 \cup F_2 \cup \dots \cup F_n) = m(F_1) + m(F_2) + \dots + m(F_n).$$

$$m^*(A) = m^*(A \cap E_1) + m^*(A \cap E_1^c \cap E_2) + m^*(A \cap E_1^c \cap E_2^c). \quad (10)$$

More generally, if we let  $A$  be arbitrary and take  $E_1 = F_1$ ,  $E_2 = F_2$  in (10) we get

$$m^*(A) = m^*(A \cap F_1) + m^*(A \cap F_2) + m^*(A \cap (F_1 \cup F_2)^c).$$

If  $F_3 \in \mathcal{M}$  is disjoint from  $F_1$  and  $F_2$  we may apply (8) with  $A$  replaced by  $A \cap (F_1 \cup F_2)^c$  and  $E$  by  $F_3$  to get

$$m^*(A \cap (F_1 \cup F_2)^c) = m^*(A \cap F_3) + m^*(A \cap (F_1 \cup F_2 \cup F_3)^c),$$

since

$$(F_1 \cup F_2)^c \cap F_3^c = F_1^c \cap F_2^c \cap F_3^c = (F_1 \cup F_2 \cup F_3)^c.$$

Substituting this back into the preceding equation gives

$$m^*(A) = m^*(A \cap F_1) + m^*(A \cap F_2) + m^*(A \cap F_3) + m^*(A \cap (F_1 \cup F_2 \cup F_3)^c).$$

$$m^*(A) = m^*(A \cap F_1) + m^*(A \cap F_2) + m^*(A \cap F_3) + m^*(A \cap (F_1 \cup F_2 \cup F_3)^c).$$

Proceeding inductively, we conclude that if  $F_1, \dots, F_n$  are pairwise disjoint elements of  $\mathcal{M}$  then

$$m^*(A) = \sum_1^n m^*(A \cap F_i) + m^*(A \cap (F_1 \cup \dots \cup F_n)^c). \quad (11)$$

Now suppose that we have a countable family  $\{F_i\}$  of pairwise disjoint sets belonging to  $\mathcal{M}$ . Since

$$\left(\bigcup_{i=1}^n F_i\right)^c \supset \left(\bigcup_{i=1}^{\infty} F_i\right)^c$$

we conclude from (11) that

$$m^*(A) \geq \sum_1^n m^*(A \cap F_i) + m^*\left(A \cap \left(\bigcup_{i=1}^{\infty} F_i\right)^c\right)$$

and hence passing to the limit

$$m^*(A) \geq \sum_1^{\infty} m^*(A \cap F_i) + m^*\left(A \cap \left(\bigcup_{i=1}^{\infty} F_i\right)^c\right).$$

Now given any collection of sets  $B_k$  we can find intervals  $\{I_{k,j}\}$  with

$$B_k \subset \bigcup_j I_{k,j}$$

and

$$m^*(B_k) \leq \sum_j \ell(I_{k,j}) + \frac{\epsilon}{2^k}.$$

So

$$\bigcup_k B_k \subset \bigcup_{k,j} I_{k,j}$$

and hence

$$m^*\left(\bigcup B_k\right) \leq \sum m^*(B_k),$$

the inequality being trivially true if the sum on the right is infinite.



So  $\sum_{i=1}^{\infty} m^*(A \cap F_k) \geq m^*(A \cap (\bigcup_{i=1}^{\infty} F_i))$ . Thus

$$\begin{aligned} m^*(A) &\geq \sum_1^{\infty} m^*(A \cap F_i) + m^*\left(A \cap \left(\bigcup_{i=1}^{\infty} F_i\right)^c\right) \geq \\ &\geq m^*\left(A \cap \left(\bigcup_{i=1}^{\infty} F_i\right)\right) + m^*\left(A \cap \left(\bigcup_{i=1}^{\infty} F_i\right)^c\right). \end{aligned}$$

The extreme right of this inequality is the left hand side of (9) applied to

$$E = \bigcup_i F_i,$$

and so  $E \in \mathcal{M}$  and the preceding string of inequalities must be equalities since the middle is trapped between both sides which must be equal.

Hence we have proved that if  $F_n$  is a disjoint countable family of sets belonging to  $\mathcal{M}$  then their union belongs to  $\mathcal{M}$  and

$$m^*(A) = \sum_i m^*(A \cap F_i) + m^*\left(A \cap \left(\bigcup_{i=1}^{\infty} F_i\right)^c\right). \quad (12)$$

If we take  $A = \bigcup F_i$  we conclude that

$$m(F) = \sum_{n=1}^{\infty} m(F_n) \quad (13)$$

if the  $F_j$  are disjoint and

$$F = \bigcup F_j.$$

So we have reproved the last assertion of the theorem using Caratheodory's definition.

Proof of 3: If  $E_n \in \mathcal{M}$  for  $n = 1, 2, 3, \dots$  then  $\bigcup_n E_n \in \mathcal{M}$ .

For the proof of this third assertion, we need only observe that a countable union of sets in  $\mathcal{M}$  can be always written as a countable disjoint union of sets in  $\mathcal{M}$ . Indeed, set

$$F_1 := E_1, \quad F_2 := E_2 \setminus E_1 = E_2 \cap E_1^c$$

$$F_3 := E_3 \setminus (E_1 \cup E_2)$$

etc. The right hand sides all belong to  $\mathcal{M}$  since  $\mathcal{M}$  is closed under taking complements and finite unions and hence intersections, and

$$\bigcup_j F_j = \bigcup E_j.$$

We have completed the proof of the theorem.

## Some consequences - symmetric differences.

The symmetric difference between two sets is the set of points belonging to one or the other but not both:

$$A \Delta B := (A \setminus B) \cup (B \setminus A).$$

Proposition.

If  $A \in \mathcal{M}$  and  $m(A \Delta B) = 0$  then  $B \in \mathcal{M}$  and  $m(A) = m(B)$ .

## Proof.

By assumption  $A \setminus B$  has measure zero (and hence is measurable) since it is contained in the set  $A \Delta B$  which is assumed to have measure zero. Similarly for  $B \setminus A$ . Also  $(A \cap B) \in \mathcal{M}$  since

$$A \cap B = A \setminus (A \setminus B).$$

Thus

$$B = (A \cap B) \cup (B \setminus A) \in \mathcal{M}.$$

Since  $B \setminus A$  and  $A \cap B$  are disjoint, we have

$$m(B) = m(A \cap B) + m(B \setminus A) = m(A \cap B) = m(A \cap B) + m(A \setminus B) = m(A).$$



## More consequences: increasing limits.

### Proposition.

Suppose that  $A_n \in \mathcal{M}$  and  $A_n \subset A_{n+1}$  for  $n = 1, 2, \dots$ . Then

$$m\left(\bigcup A_n\right) = \lim_{n \rightarrow \infty} m(A_n).$$

If  $A := \bigcup A_n$  we write the hypotheses of the proposition as  $A_n \uparrow A$ . In this language the proposition asserts that

$$A_n \uparrow A \Rightarrow m(A_n) \rightarrow m(A).$$

## Proof.

Setting  $B_n := A_n \setminus A_{n-1}$  (with  $B_1 = A_1$ ) the  $B_i$  are pairwise disjoint and have the same union as the  $A_i$  so

$$\begin{aligned} m\left(\bigcup A_n\right) &= \sum_{i=1}^{\infty} m(B_i) = \lim_{n \rightarrow \infty} \sum_{i=1}^n m(B_i) \\ &= \lim_{n \rightarrow \infty} m\left(\bigcup_{i=1}^n B_i\right) = \lim_{n \rightarrow \infty} m(A_n). \end{aligned}$$



## More consequences: decreasing limits.

### Proposition.

If  $C_n \supset C_{n+1}$  is a decreasing family of sets in  $\mathcal{M}$  and  $m(C_1) < \infty$  then

$$m\left(\bigcap C_n\right) = \lim_{n \rightarrow \infty} m(C_n).$$

### Proof.

Let  $C = \bigcap C_n$  and  $A_n := C_1 \setminus C_n$ . Then  $A_n$  is a monotone increasing family of sets as in the preceding proposition so  $m(C_1) - m(C_n) = m(A_n) \rightarrow m(C_1 \setminus C) = m(C_1) - m(C)$ . Now subtract  $m(C_1)$  from both sides. □

Taking  $C_n = [n, \infty)$  shows that we need  $m(C_k) < \infty$  for some  $k$ .



## Constantin Carathéodory



**Born: 13 Sept 1873 in Berlin, Germany**

**Died: 2 Feb 1950 in Munich, Germany**

# Axiomatic approach.

We will now take the items in Theorem 6 as axioms: Let  $X$  be a set. (Usually  $X$  will be a topological space or even a metric space). A collection  $\mathcal{F}$  of subsets of  $X$  is called a  $\sigma$  field if:

- $X \in \mathcal{F}$ ,
- If  $E \in \mathcal{F}$  then  $E^c = X \setminus E \in \mathcal{F}$ , and
- If  $\{E_n\}$  is a sequence of elements in  $\mathcal{F}$  then  $\bigcup_n E_n \in \mathcal{F}$ ,

The intersection of any family of  $\sigma$ -fields is again a  $\sigma$ -field, and hence given any collection  $\mathcal{C}$  of subsets of  $X$ , there is a smallest  $\sigma$ -field  $\mathcal{F}$  which contains it. Then  $\mathcal{F}$  is called the  $\sigma$ -field **generated** by  $\mathcal{C}$ .

# Borel sets.

If  $X$  is a metric space, the  $\sigma$ -field generated by the collection of open sets is called the **Borel**  $\sigma$ -field, usually denoted by  $\mathcal{B}$  or  $\mathcal{B}(X)$  and a set belonging to  $\mathcal{B}$  is called a **Borel set**.

# Measures on a $\sigma$ -field.

Given a  $\sigma$ -field  $\mathcal{F}$  a (non-negative) **measure** is a function

$$m : \mathcal{F} \rightarrow [0, \infty]$$

such that

- $m(\emptyset) = 0$  and
- **Countable additivity:** If  $F_n$  is a disjoint collection of sets in  $\mathcal{F}$  then

$$m\left(\bigcup_n F_n\right) = \sum_n m(F_n).$$

In the countable additivity condition it is understood that both sides might be infinite.

# Outer measures.

An **outer measure** on a set  $X$  is a map  $m^*$  to  $[0, \infty]$  defined on the collection of *all* subsets of  $X$  which satisfies

- $m^*(\emptyset) = 0$ ,
- **Monotonicity:** If  $A \subset B$  then  $m^*(A) \leq m^*(B)$ , and
- **Countable subadditivity:**  $m^*(\bigcup_n A_n) \leq \sum_n m^*(A_n)$ .

# Measures from outer measures via Caratheodory.

Given an outer measure,  $m^*$ , we defined a set  $E$  to be **measurable** (relative to  $m^*$ ) if

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c)$$

for all sets  $A$ . Then Caratheodory's theorem that we proved just now asserts that the collection of measurable sets is a  $\sigma$ -field, and the restriction of  $m^*$  to the collection of measurable sets is a measure which we shall usually denote by  $m$ .

# Terminological conflict.

There is an unfortunate disagreement in terminology, in that many of the professionals, especially in geometric measure theory, use the term “measure” for what we have been calling “outer measure”. However we will follow the above conventions which used to be the old fashioned standard.

## The next task.

An obvious task, given Caratheodory's theorem, is to look for ways of constructing outer measures.

This will be the subject of the next lecture.



Over the next few slides I want to give remarkable applications of our two “monotone convergence” propositions:

$$A_n \uparrow A \Rightarrow m(A_n) \rightarrow m(A)$$

and (with the obvious notation)

$$C_n \downarrow C \text{ and } m(C_1) < \infty \Rightarrow m(C_n) \rightarrow m(C).$$

## Some notation

Let  $E_n$  be a sequence of measurable sets. The following definitions of the set  $E$  are synonymous:

$$\begin{aligned}
 E &:= \{E_n \text{ i.o.}\} = \{E_n \text{ infinitely often}\} \\
 &:= \limsup E_n \\
 &:= \bigcap_{k=1}^{\infty} \bigcup_{n \geq k} E_n \\
 &:= \{x \mid \forall k \exists n(k) \geq k \text{ such that } x \in E_{n(k)}\} \\
 &:= \{x \mid x \in E_n \text{ for infinitely many } E_n\}.
 \end{aligned}$$

# The first Borel-Cantelli lemma (BC1)

## Lemma

*Suppose that  $\sum m(E_n) < \infty$  then  $m(\limsup E_n) = 0$ .*

## Proof.

Let  $G_k = \bigcup_{n \geq k} E_n$ . So  $\bigcup_{k \leq n \leq p} E_n \uparrow G_k$ . But

$$m\left(\bigcup_{k \leq n \leq p} E_n\right) \leq \sum_{k \leq n \leq p} m(E_n)$$

and so  $m(G_k) \leq \sum_k^\infty m(E_n)$ . Since  $\limsup E_n \subset G_k$  for all  $k$  and  $\sum m(E_n) < \infty$  we conclude that  $m(\limsup E_n)$  is less than any positive number by choosing  $k$  sufficiently large. □

## Some probabilistic language

We now restrict to the case where  $m(X) = 1$  and use  $\mathbb{P}$  instead of  $m$ . A measurable set is now called an **event**. A sequence of events  $E_n$  is called **independent** if, for every  $k \in \mathbb{N}$ , if all the  $i_1, \dots, i_k$  are distinct then

$$\mathbb{P}(E_{i_1} \cap \dots \cap E_{i_k}) = \prod \mathbb{P}(E_{i_j}).$$

# The second Borel-Cantelli lemma (BC2)

## Lemma

Let  $E_n$  be a sequence of independent events. Then  
 $\sum \mathbb{P}(E_n) = \infty \Rightarrow \mathbb{P}(E_n \text{ i.o.}) = 1.$

Notice that  $E_n \text{ i.o.} = \bigcap_k \bigcup_{n \geq k} E_n$  so its complement is

$$\bigcup_k \bigcap_{n \geq k} E_n^c.$$

We must show that this has probability zero, and for this it is enough to show that  $\bigcap_{n \geq k} E_n^c$  has probability zero. Now

$$\bigcap_{k \leq n \leq \ell} E_n^c \downarrow \bigcap_{n \geq k} E_n^c$$

so it is enough to show that  $\mathbb{P}\left(\bigcap_{k \leq n \leq \ell} E_n^c\right) \rightarrow 0$  as  $\ell \rightarrow \infty$ .

## Proof.

Let  $p_n := \mathbb{P}(E_n)$  so that  $\mathbb{P}(E_n^c) = 1 - p_n$ . Then, by independence,

$$\mathbb{P}\left(\bigcap_{k \leq n \leq \ell} E_n^c\right) = \prod_{k \leq n \leq \ell} (1 - p_n).$$

Now for  $x \geq 0$ ,  $1 - x \leq e^{-x}$  so

$$\prod_{k \leq n \leq \ell} (1 - p_n) \leq e^{-\sum_{k \leq n \leq \ell} p_n}$$

and since  $\sum p_n = \infty$  this last expression  $\rightarrow 0$ .



Notice we get a “0 or 1 law” in that for a sequence  $E_n$  of independent events we conclude that  $\mathbb{P}(E_n \text{ infinitely often})$  is either zero or one.