FUNDAMENTALS OF REAL ANALYSIS
by

Doğan Çömez

II. MEASURES AND MEASURE SPACES

II.1. Prelude

Recall that the Riemann integral of a real-valued function f on an interval $[a, b]$ is defined as

$$\int_a^b f(x)dx := \lim_{\|\mathcal{P}\|\to 0} \sum_{i=1}^{n} f(\overline{x_i})(x_i - x_{i-1}), \text{ (if the limit exists)}$$

where $\mathcal{P} = \{a = x_1 < x_2 < \cdots < x_{n-1} < x_n = b\}$ is a partitioning of $[a, b]$, $\|\mathcal{P}\| = \max_i |x_i - x_{i-1}|$, and $\overline{x_i} \in [x_{i-1}, x_i]$ is arbitrary. Then, it follows that if $f : [a, b] \to \mathbb{R}$ is a continuous function, then it is Riemann-integrable. Also, if $f : [a, b] \to \mathbb{R}$ has finitely many bounded jump discontinuities, then it is Riemann-integrable. Indeed, there are functions with infinitely many bounded jumps which are still Riemann-integrable. However, there are "very simple" functions with countably many bounded jumps which are not Riemann-integrable. For example, let

$$f(x) = \begin{cases} 1 \text{ if } x \in [0, 1] \cap \mathbb{Q} \\ 0 \text{ if } x \in [0, 1] \cap \mathbb{Q}^c. \end{cases}$$

Then, $\limsup_{\|\mathcal{P}\|\to 0} \sum_{i=1}^{n} f(\overline{x_i})(x_i - x_{i-1}) = 1 \neq 0 = \liminf_{\|\mathcal{P}\|\to 0} \sum_{i=1}^{n} f(\overline{x_i})(x_i - x_{i-1})$!

Riemann integral also suffers from a convergence problem: let $\{r_i\}$ be an enumeration (listing) of rational numbers in $[0, 1]$, and define a sequence of functions $\{f_n\}$ on $[0, 1]$ by

$$f_n(x) = \begin{cases} 1 \text{ if } x = r_1,\ r_2,\ r_3,\ \ldots,\ r_n \\ 0 \text{ otherwise,} \end{cases}$$

and f be the function in the example above. Then each f_n is Riemann-integrable, $f_n \to f$ pointwise on $[0, 1]$, but f is not Riemann-integrable.

Verdict: Riemann integral has several deficits!

Next, observe that if f is a positive, continuous function on $[a, b]$ or has finite jump discontinuities, then $\int_a^b f dx$ gives the area of the region $R = \{(x, y) : a \leq x \leq b,\ 0 \leq y \leq f(x)\}$. However, if the discontinuities of f is more "complicated" or if f is continuous but has oscillations on a "large" set, then f need not be Riemann-integrable; hence, for such regions "area" is not defined.

Question. Can we define area of a region without involving Riemann-integral? If so, how?

This was the question asked by mathematicians of late 19th century that led to the concept of "measure" that would measure the size (length, area, volume, . . . etc.) of a set (in $\mathbb{R}^n$). For, one needs a working definition of *measure.*

Recall that the area of a region $A \subset \mathbb{R}^2$ bounded by two "nice" (continuous, piecewise linear) functions as $A = \{(x, y) : a \leq x \leq b,\ f(x) \leq y \leq g(x)\}$ is given by the Riemann-integral $\int_a^b |f(x) - g(x)|dx$. This definition is still valid if f and g have finitely many bounded jump discontinuities. This is a good start; however, as seen in the example above, there are functions with countably many bounded jump discontinuities which are not Riemann-integrable.

Furthermore, there are continuous curves (so called "space-filling curves") which do not fit to this scheme. Therefore, defining area via Riemann-integral has serious inadequacies.

Here comes Peano. In order to address this deficiency, Peano proposed the following geometric approach (inspired from Archimedes' idea on calculating areas enclosed by conic sections). For $S \subset \mathbb{R}^2$, define

$$a_i(S) = \text{inner area of } S = \text{supremum of areas of all polygons contained in S},$$
$$a_o(S) = \text{outer area of } S = \text{infimum of areas of all polygons containing S},$$

and define $A(S) = a_i(S)$ if both the inner and outer area exist and the equality $a_o(S) = a_i(S)$ holds. Although this is a significant improvement over the definition of area via Riemann integral, it still has some drawbacks. For instance, if $S = \{(x, y) : x, y \in [0, 1] \cap \mathbb{Q}\}$, then $a_o(S) = 1 \neq 0 = a_i(S)$. Thus, some rather simple regions do not have area!

Jordan's try. Inspired by Peano, Jordan proposed using the concept of "content" to address the problem. For $S \subset [a, b]$, let $\mathcal{P} = \{(x_i, x_{i+1})\}$ be any finite partition of $[a, b]$ by open intervals. Let

$$J_i(S) = \{\sum l(I_i) : S \supset I_i \in \mathcal{P} \text{ contains interior points of } S\},$$
$$J_o(S) = \{\sum l(I_i) : I_i \in \mathcal{P} \text{ contains points of } S \cup \partial S\},$$

and let $c_i(S) = \sup_{\mathcal{P}} J_i(S)$ and $c_i(S) = \inf_{\mathcal{P}} J_o(S)$. Then, if $c_i(S)$ and $c_o(S)$ exist and are equal, define $c(S) =$ **content of** $S = c_i(S)$ and call S *Jordan-measurable.* It turns out that form many sets in $\mathbb{R}$ their content is well defined (which includes regions defined by space filling curves). Jordan also defined integral of a real-valued function f on $[a, b]$ as follows. Let

$$\int_{\underline{a}}^{b} f = \sup_{\mathcal{P}} \sum \inf_{E_i} f(x).c(E_i)$$
$$\int_{a}^{\bar{b}} f = \inf_{\mathcal{P}} \sum \sup_{E_i} f(x).c(E_i),$$

where $\mathcal{P} = \{E_i\}$ is a partitioning of $[a, b]$ into Jordan measurable sets. If $\int_{\underline{a}}^{b} f = \int_{a}^{\bar{b}} f$, then the common value is called the integral of f.

Jordan integral improves Riemann integral; furthermore, the collection of Jordan-measurable sets is richer than those having area in the sense of Peano. On the other hand, the content concept also has some problems; for example, $c_i([0,1] \cap \mathbb{Q}) = 0 = c_i([0,1] \cap \mathbb{Q}^c)$ and $c_o([0,1] \cap \mathbb{Q}) = 1 = c_o([0,1] \cap \mathbb{Q}^c)$; thus, both of these sets are not Jordan-measurable. Furthermore, there are even some open (or closed) bounded sets which are not Jordan-measurable, which is a serious drawback.

Despite it's shortcomings, Jordan measure (content) and Jordan measurable sets have the following "nice" properties:

1) If A and B are Jordan measurable, then so are $A \cap B$, $A \cup B$, $A \setminus B$, $B \setminus A$ (closure under set operations).
2) $c(E) \geq 0$ (non-negativity).
3) If E and F are disjoint Jordan measurable sets, then $c(E \cup F) = c(E) + c(F)$ (finite additivity); if E and F are not necessarily disjoint, $c(E \cup F) \leq c(E) + c(F)$ (subadditivity).
4) If $A \subset B$ are Jordan measurable sets, then $c(A) \leq c(B)$ (monotonicity).
5) If A is Jordan measurable and α is a real number, then $\alpha + A$ is also Jordan measurable and $c(A) = c(\alpha + A)$ (translation invariance).

Exercise Prove the properties 2)-5) stated above.

Borel's approach. The drawback of Jordan-measurability led Borel to define the properties the measure of sets should satisfy explicitly. He stated that:

1) measure of a set should be non-negative
2) measure of a (finite or infinite) union of disjoint sets should be the sum of measures of the individual sets
3) if $A \subset B$, then measure of $B \setminus A$ should be measure of B minus measure of A
4) every set whose measure is non-zero should be uncountable, and
5) the location of the set should not affect the measure of the set (i.e. measure of A and measure of $\alpha + A$ should be the same).

Notice that, these are essentially the features of Jordan content. Having described properties that a measure must carry, Borel also restricted it to those sets which are *constructible*, where he defined a set constructible if it is

a) a union (finite or countably infinite) of disjoint intervals,
b) the complement of a constructible set A wrt any other constructible set B covering A.

He prescribed that the measure of an interval would be its length. With these, he hoped that: (i) any subset of $\mathbb{R}$ would be constructible, and (ii) a measure satisfying the properties listed above could be constructed.

Bad news! (Or, may be good!?) It turns out that **it is impossible to construct a measure on subsets of $\mathbb{R}$ having all five properties above** (see pp: 19-20 in the text). However, when $n = 1, 2$, it is possible to construct a set function on subsets of $\mathbb{R}^n$ if only *finite additivity* is required (Banach), and when $n \geq 3$ even this is impossible (Hausdorff). On the other hand, constructing a set function on subsets of $\mathbb{R}^n$ satisfying first four properties is possible (though very delicate).

The property (5) above is, for obvious reasons (it makes sense!), very essential; hence, having a set function satisfying only first four properties does not address the problem. Furthermore, it turns out that **not every subset of $\mathbb{R}$ is constructible** either (we'll see an example of such a set later). Thus, since we would like to preserve as much of the properties (1)-(5) and since many sets constructed over intervals seem to satisfy them, we have one option to follow: weaken one of the properties/requirements of Borel. Weakening (1), (3) and (4) does not make sense. As seen above, although weakening (4) by requiring **only** finitely additivity does not solve the problem, one can weaken (2) in an appropriate fashion while keeping measure well-defined. These considerations provide compelling rationale to give up defining a measure on all of $\mathcal{P}(\mathbb{R})$, and, instead, let a measure be defined on certain (proper) subclasses of $\mathcal{P}(\mathbb{R})$.

First of all, we must be careful in applying "additivity" of a measure m. If additivity also includes adding over not-necessarily countable index sets, then we would encounter paradoxical situations. For instance, $(0, 1] = \cup_{x\in(0,1]}\{x\}$; hence,

$$1 = m((0, 1]) = \sum_{x\in(0,1]} m(\{x\}) = \sum_{x\in(0,1]} 0 = 0!$$

Therefore, it'd be safer to restrict it to "countable additivity". This clears way to

Program. Develop: (1) a good understanding of constructible sets, and (2) a good concept of measure on constructible sets.

Let's start with (1). Following Borel's prescription of constructibility, we see that constructible sets must satisfy the following axioms:

(i) if A and B are constructible, then $A \cup B$ is also constructible
(i′) if $\{A_i\}$ is a countable collection of constructible sets, then $\cup_i A_i$ is also constructible
(ii) if A is constructible, then so is A^c
(iii) $\mathbb{R}$ and $\emptyset$ are constructible.

Notice that, since intervals are *a priori* constructible, all open (closed) sets are constructible (why?), and so are arbitrary union (intersection) of closed (open) sets.

Definition. The smallest collection $\mathcal{B}$ of all subsets of $\mathbb{R}$ which contains all open sets and satisfying (i′), (ii) and (iii) is called *Borel sets* (or *Borel σ-algebra*).

Remark. From the definition, the collection $\mathcal{B}$ includes:

(a) all closed sets
(b) all arbitrary union of closed sets ($\mathcal{F}_\sigma$ sets)
(c) all arbitrary intersection of open sets ($\mathcal{G}_\delta$ sets)
(d) all arbitrary intersection of sets in (b) ($\mathcal{F}_{\sigma\delta}$ sets)
(e) all arbitrary union of sets in (c) ($\mathcal{G}_{\delta\sigma}$ sets)

and so on. Notice that this process goes on indefinitely; hence the definition of $\mathcal{B}$ is stated above in a “peculiar” manner.

Exercises. 1. Show that $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{Q}^c$, $\mathcal{C}$ are Borel sets, where $\mathcal{C}$ is the Cantor set. Which one of the classes $\mathcal{F}_\sigma$, $\mathcal{G}_\delta$, $\mathcal{F}_{\sigma\delta}, \ldots$ each of these sets belong?

2. Prove that translate of any $\mathcal{F}_\sigma$-set ($\mathcal{G}_\delta$-set) is a $\mathcal{F}_\sigma$-set ($\mathcal{G}_\delta$-set)

Question. Does the collection $\mathcal{B}$ actually exist?

The answer is a resounding **Yes** as the following statement proves.

Fact.1 Given any collection $\mathcal{O}$ of subsets of $\mathbb{R}$, there exists a smallest collection $\mathcal{A}$ of subsets of $\mathbb{R}$ satisfying (i′), (ii), (iii) and containing $\mathcal{O}$.

Note. Such a family $\mathcal{A}$ is called a *σ-algebra* (of subsets of $\mathbb{R}$).

Proof. (of Fact.1) Let Δ be the family of all σ-algebras of subsets of $\mathbb{R}$ that contain $\mathcal{O}$. Then Δ is nonempty (it contains $P(\mathbb{R})$). Let $\mathcal{A} = \cap\{\mathcal{D} : \mathcal{D} \in \Delta\}$. Then, $\mathcal{O} \subset \mathcal{A}$. Furthermore,

(i) if $A, B \in \mathcal{A}$, then $A, B \in \mathcal{D}$ for all $\mathcal{D} \in \Delta$. So, since such a $\mathcal{D}$ is a σ-algebra, $A \cup B \in \mathcal{D}$ for all $\mathcal{D} \in \Delta$. Hence, $A \cup B \in \mathcal{A}$.

(ii) if $A \in \mathcal{A}$, then $A \in \mathcal{D}$ for all $\mathcal{D} \in \Delta$; and it follows that $A^c \in \mathcal{D}$ for all $\mathcal{D} \in \Delta$. Hence, $A^c \in \mathcal{A}$.

(iii) By definition, $\emptyset, \mathbb{R} \in \mathcal{D}$ for all $\mathcal{D} \in \Delta$; and hence, $\emptyset, \mathbb{R} \in \mathcal{A}$.

Therefore, $\mathcal{A}$ is a σ-algebra. From the definition of $\mathcal{A}$, if $\mathcal{D}$ is a σ-algebra containing $\mathcal{O}$, then $\mathcal{A} \subset \mathcal{D}$; hence, $\mathcal{A} \subset \mathcal{D}$ for all $\mathcal{D} \in \Delta$. Hence, $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{O}$. ∎

Remark. $|\mathcal{B}| = c$. (See Real and Abstract Analysis, by Hewitt & Stromberg, Theorem 10.23, pp:133-134.) It is well known that $|\mathcal{P}(\mathbb{R})| = 2^c > c$; hence, $\mathcal{B}$ is a **proper** subset of $\mathcal{P}(\mathbb{R})$! Thus, it appears that the smallest collection of sets that are constructible over all intervals in $\mathbb{R}$ does not include **all** subsets of $\mathbb{R}$.

This remark suggests that in order to complete the program identified above, we might keep (1), (3)-(5) intact and restrict (2) to **countable additivity** and have measures defined on largest possible subcollection (σ-algebra) $\mathcal{A}$ **smaller** than $\mathcal{P}(\mathbb{R})$.

Now, let's see how can we approach the problem of constructing a measure m on a σ-algebra. Following Borel, the starting point is defining measure on finite sets and intervals. Define measure of an interval as its length. So $m([a,b]) = l([a,b]) = l((a,b)) = l((a,b]) = l([a,b)) = b-a = m((a,b))$. If $a = -\infty$ or $b = \infty$, then $m([a,b]) = l([a,b] = \infty$. Since $[a,b] = [a,b) \cup \{b\}$, we must have $m(\{b\}) = 0$ for any singleton $\{b\} \subset \mathbb{R}$; consequently, measure of any finite set is 0. Also, $m([a,b]) = l([a,b]) < l([a,2b]) = m([a,2b])$; hence $m(A) \leq m(B)$ should hold if $A \subset B$. This discussion, with a little luck, tells us that if we are to construct a measure on subsets of $\mathbb{R}$, on that collection we'd hope it to satisfy the following properties:

(i) $0 \leq m(A) \leq \infty$
(ii) $m(A) \leq m(B)$ if $A \subset B$ (monotonicity)
(iii) $m(\emptyset) = 0 = m(\{a\}),\ \forall a \in \mathbb{R}$
(iv) $m(I) = l(I)$ for any interval $I \subset \mathbb{R}$
(v) $m(\cup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} m(A_k)$ for all disjoint collection of sets $\{A_k\}$ (countable additivity)
(vi) $m(a+A) = m(A),\quad \forall a \in \mathbb{R}$ (translation invariance).

Exercises

1. Recall that for $S \subset [a,b]$ and $\mathcal{P} = \{(x_i, x_{i+1})\}$ any finite partition of $[a,b]$ by open intervals, we define

$$J_i(S) = \{\sum l(I_i) : I_i \subset S,\ I_i \in \mathcal{P} \text{ contains interior points of } S\},$$

$$J_o(S) = \{\sum l(I_i) : I_i \in \mathcal{P} \text{ contains points of } S \cup \partial S\},$$

and let $c_i(S) = \sup_{\mathcal{P}} J_i(S)$ and $c_o(S) = \inf_{\mathcal{P}} J_o(S)$. Then, if $c_i(S)$ and $c_o(S)$ exist and are equal, we define $c(S) =$ **content of** $S = c_i(S)$ and call S *Jordan-measurable.* Prove that

(i) $c(E) \geq 0$ (non-negativity).

(ii) If $A \subset B$ are Jordan-measurable sets, then $c(A) \leq c(B)$ (monotonicity).

(iii) If E and F are disjoint Jordan-measurable sets, $c(E \cup F) = c(E) + c(F)$ (finite additivity); if E and F are not necessarily disjoint, $c(E \cup F) \leq c(E) + c(F)$ (subadditivity).

(iv) If A is Jordan measurable and α is a real number, then $\alpha + A$ is also Jordan measurable and $c(A) = c(\alpha + A)$ (translation invariance).

2. Provide simple justifications that any $\mathcal{F}_\sigma$, $\mathcal{G}_\delta$, $\mathcal{F}_{\sigma\delta}$ and $\mathcal{G}_{\delta\sigma}$-set is a Borel set.

3. (i) Show that $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{Q}^c$, $\mathcal{C}$ are Borel sets, where $\mathcal{C}$ is the Cantor set. Which one of the classes $\mathcal{F}_\sigma$, $\mathcal{G}_\delta$, $\mathcal{F}_{\sigma\delta}, \ldots$ each of these sets belong?

(ii) Prove that translate of any $\mathcal{F}_\sigma$-set ($\mathcal{G}_\delta$-set) is a $\mathcal{F}_\sigma$-set ($\mathcal{G}_\delta$-set).

II.2. Lebesgue Measure

Having Jordan's improvement on Peano's original ideas and Borel's specifications of a measure and constructible sets, Lebesgue developed a theory of measure that addressed the limitations of Jordan measure and applicable to a large subclass of $\mathcal{P}(\mathbb{R})$ that avoids pathological sets mentioned above. In this section, we will construct this measure known as **Lebesgue measure.** As expected, we'll begin with measure of intervals.

Clearly, we have the following for bounded intervals:

a) $0 \leq l(I) < \infty$
b) $l(I) \leq l(J)$ if $I \subset J$ (monotonicity)
c) $l(I \cup J) = l(I) + l(J)$ if $I,\ J$ are disjoint (finite additivity)
d) $l(a + I) = l(I), \quad \forall a \in \mathbb{R}$ (translation invariance).

Exercise. Let I be a bounded open interval and $I_1, I_2, \ldots, I_n$ be a finite collection of bounded open intervals. If $I \subset \cup_{i=1}^{n} I_i$, then $\ell(I) \leq \sum_{i=1}^{n} \ell(I_i)$.

Fact.2 Let I be a bounded open interval and $I_1, I_2, \ldots, I_n, \ldots$ be a countable collection of bounded open intervals such that $I \subset \cup_{i=1}^{\infty} I_i$. Then $\ell(I) \leq \sum_{i=1}^{\infty} \ell(I_i)$.

Proof. Let $I = (a, b)$ and let $\epsilon > 0$ be arbitrary. Consider

$$I' = (a - \frac{\epsilon}{4}, a + \frac{\epsilon}{4}) \text{ and } I'' = (b - \frac{\epsilon}{4}, b + \frac{\epsilon}{4}).$$

Then the collection $\{I_i\}_{i=1}^{\infty} \cup \{I', I''\}$ is an open cover for $[a, b]$. Hence by Heine-Borel theorem, there exists a finite sub-cover, say $\{I_i\}_{i=1}^{\infty} \cup \{I', I''\}$. Now, apply the Exercise above. ∎

Remarks. 1. If the collection $I_1, I_2, \ldots, I_n, \ldots$ in Fact.2 above were a disjoint, countable collection of bounded open intervals covering I, then $\ell(I) \leq \sum_{i=1}^{\infty} \ell(I_i)$ would follow (almost) trivially from (c) above. (Exercise)

2. If, furthermore, $\{I_i\}$ is a disjoint, countable collection of bounded open intervals with $I = \cup_{i=1}^{\infty} I_i$, then $\ell(I) = \sum_{i=1}^{\infty} \ell(I_i)$. (Exercise)

Now, we would like to extend these (nice) properties of the measure $m(= \ell)$ on intervals to more general sets of real numbers.

Definition. The set function $m^* : \mathcal{P}(\mathbb{R}) \to \mathbb{R}_+^{\#}$ defined by

$$m^*(A) = \inf \left\{ \sum_{k=1}^{\infty} \ell(I_k) : A \subseteq \bigcup_{k=1}^{\infty} I_k, \text{ each } I_k \text{ is a bounded open interval} \right\}$$

is called the **(Lebesgue) outer measure**.

Remarks. It is easy to show that:

(i) $m^*(\emptyset) = 0$ and, $m^*(\{a\}) = 0$ for all $a \in \mathbb{R}$.
(ii) $m^*(A) \leq m^*(B)$ if $A \subset B$.

Exercises. 1. Calculate $m^*(A)$ if (i) A is a singleton or a finite set, (ii) $A = \mathbb{Q} \cap [0, 1]$.

2) Show that $m^*(A) = m^*(\alpha + A)$ for all $\alpha \in \mathbb{R}$.

Fact.3 For any interval $I \subset \mathbb{R}, \quad m^*(I) = \ell(I)$.

Proof. Assume $I = (a, b)$, then , given any $\epsilon > 0$, we have $I \subset (a - \frac{\epsilon}{2}, b + \frac{\epsilon}{2})$. Then, by Remark (ii) above, $m^*(I) \leq \ell(I) < b - a + \epsilon$ which implies that $m^*(I) \leq b - a = \ell(I)$. On the other hand, $b - a$ is a lower bound for the set of numbers

$$U = \left\{ \sum_{i=1}^{\infty} \ell(I_i) : I \subset \bigcup_{i=1}^{\infty} I_i,\ \{I_i\} \text{ open} \right\}.$$

Since $m^*(I) = \inf(U)$, which must be greater than or equal to any other lower bound, we have $m^*(I) \geq b - a = \ell(I)$. Hence $m^*(E) = \ell(I)$.

Next assume that $I = [a, b]$. Then, for all $\epsilon > 0$ $(a, b) \subset I \subset (a - \epsilon, b + \epsilon)$. Thus,

$$\begin{aligned} m^*(a,b) &\leq m^*(I) \leq m^*(a-\epsilon, b+\epsilon) \\ &\Rightarrow \ell(I) \leq m^*(I) \leq \ell(E) + 2\varepsilon \\ &\Rightarrow \ell(I) = m^*(I). \end{aligned}$$

The other two bounded cases are proved similarly.

If I is an unbounded interval (i.e., if $a = -\infty$ or $b = \infty$), then there exists an interval $I' \subset I$ with $\ell(I') = n$. Then $m^*(I) \geq m^*(I') = n$ for all $n \geq 1$; hence, $m^*(I) = \infty = \ell(I)$. ∎

Fact.4 The outer measure m^* is countably sub-additive, namely, if $A_1, A_2, \ldots$ is a sequence of subsets of $\mathbb{R}$, then

$$m^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} m^*(A_i).$$

Proof. Fix $\epsilon > 0$. Given any A_i, pick a cover of A_i consisting of open intervals $\{I_{i,n}\}$ such that

$$m^*(A_i) \leq \sum_{n=1}^{\infty} \ell(I_{i,n}) < m^*(A_i) + \frac{\varepsilon}{2^i}.$$

Next consider the collection $\{I_{i,n}\}_{i,n=1}^{\infty}$, which covers $\bigcup_{i=1}^{\infty} A_i$. Hence

$$m^*\left(\bigcup_{i=1}^{\infty} A_i\right) = \inf\left\{\sum_{k=1}^{\infty} \ell(J_k) : \bigcup_{k=1}^{\infty} J_k \supset \bigcup_{i=1}^{\infty} A_i\right\}.$$

Therefore,

$$m^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i,n=1}^{\infty} \ell(I_{i,n}) = \sum_{i=1}^{\infty}\sum_{n=1}^{\infty} \ell(I_{i,n}) < \sum_{i=1}^{\infty}\left(m^*(A_i) + \frac{\varepsilon}{2^i}\right) = \sum_{i=1}^{\infty} m^*(A_i) + \varepsilon. \quad \blacksquare$$

Remark. It was shown by Vitali (1905) that there are $A, B \subset \mathbb{R}$ disjoint such that $m^*(A) > 0$, $m^*(B) > 0$, and $m^*(A \cup B) \neq m^*(A) + m^*(B)$.

Fact.5 m^* is finitely additive implies m^* is countably additive.

Proof. If $\{A_i\}_{i=1}^{\infty}$ is a disjoint family of subsets of $\mathbb{R}$ and m^* is finitely additive, then by Fact.4

$$\sum_{i=1}^{n} m^*(A_i) = m^*\left(\bigcup_{i=1}^{n} A_i\right) \leq m^*\left(\bigcup_{-=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} m^*(A_i).$$

Now by letting $n \to \infty$ we have $m^*(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} m^*(A_i)$. ∎

Consequently, in order to have m^* be countably additive, we need to seek a smaller class of subsets of $\mathbb{R}$ (than $\mathcal{P}(\mathbb{R})$) on which m^* is finitely additive.

A simplification: From Fact.4 and Fact.5, in order to show that m^* is finitely additive on a class of sets, all we need to show is that $m^*(A \cup B) \geq m^*(A) + m^*(B)$ for any pair A, B in that class.

Definition. A set $E \subset \mathbb{R}$ is called **measurable** if given any $A \subset \mathbb{R}$, then

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c), \quad \text{i.e., } E \text{ splits } A \ m^*\text{-additively}.$$

Remarks. 1. $\emptyset$ and $\mathbb{R}$ are measurable.

2. E measurable implies that E^c is measurable.

Question. What is the structure and the size of the the class of measurable sets?

In order to answer this question we need to make an observation.

Fact.6 If $m^*(E) = 0$, then E is measurable.

Proof. Let $A \in \mathbb{R}$. Then $(A \cap E) \subset E$ and $(A \cap E^c) \subset A$. Hence by monotonicity of m^*

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

Hence $m^*(A \cap E) + m^*(A \cap E^c) \leq m^*(A)$. Conversely we notice that

$$m^*(A) \geq m^*(A \cap E^c) = 0 + m^*(A \cap E^c) = m^*(A \cap E) + m^*(A \cap E^c).$$

Therefore, E is measurable. ∎

Definition. A nonempty collection $\mathcal{G}$ of subsets of $\mathbb{R}$ is called an **algebra** if

i) $\mathbb{R} \in \mathcal{G}$,
ii) if $\{A_i\}_{i=1}^n \subset \mathcal{G}$, then $\cup_{i=1}^n A_i \in \mathcal{G}$, and
iii) $A \in \mathcal{G}$ implies $A^c \in \mathcal{G}$.

Remark. It follows from (i) and (iii) that $\emptyset \in \mathcal{G}$. Also, it follows from (ii) and (iii) that if $\{A_i\}_{i=1}^n \subset \mathcal{G}$, then $\cap_{i=1}^n A_i \in \mathcal{G}$.

Definition. A nonempty collection $\mathcal{G}$ of subsets of $\mathbb{R}$ is called a **σ-algebra** if it is an algebra and (ii′) if $\{A_i\}_{i=1}^\infty \subset \mathcal{G}$, then $\cup_{i=1}^\infty A_i \in \mathcal{G}$.

Theorem.1 $\mathcal{F} = \{E \subset \mathbb{R} : \ E \text{ is measurable}\}$ is a $\sigma-$algebra.

Proof. We first note that $\emptyset, \mathbb{R} \in \mathcal{F}$. Hence (i) is satisfied. Also $E \in \mathcal{F}$ which implies that $E^c \in \mathcal{F}$. Thus (iii) is also satisfied. Therefore, we must show $\{E_i\}_{i=1}^\infty \subset \mathcal{F} \Rightarrow \bigcup_{i=1}^\infty E_i \in \mathcal{F}$.

First we will consider the finite case. Let $E_1, E_2 \in \mathcal{F}$. In order to prove that $E_1 \cup E_2 \in \mathcal{F}$ we need to show for all $A \subset \mathbb{R}$ that

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

Observe that

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

Now

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

$$\begin{aligned} &\leq \ m^*(A \cap E_1) + m^*(A \cap E_2 \cap E_1^c) + m^*(A \cap E_1^c \cap E_2^c) \\ &= \ m^*(A \cap E_1) + m^*(A \cap E_1^c) = m^*(A) \end{aligned}$$

since $E_2 \in \mathcal{F}$. Hence $(E_1 \cup E_2) \in \mathcal{F}$. Therefore $\mathcal{F}$ is an algebra.

Claim. If $\{E_1, E_2, \ldots, E_n\}$ is a disjoint family of sets in $\mathcal{F}$, then for all $A \subset \mathbb{R}$ we have

$$m^*\left(A \cap \left(\bigcup_{i=1}^n E_i\right)\right) = \sum_{i=1}^n m^*(A \cap E_i).$$

(Exercise: Prove this claim. Hint: Use induction.)

Next let $\{E_i\} \subset \mathcal{F}$ be an arbitrary (countable) family of sets. Without loss of generality, we can assume that it is a disjoint family. Let $E = \bigcup_{i=1}^\infty E_i$. We now need to show for all $A \subset \mathbb{R}$

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

First observe that

$$E^c = \bigcap_{i=1}^{\infty} E_i^c \subset \bigcap_{i=1}^{n} E_i^c = \left(\bigcup_{i=1}^{n} E_i\right)^c$$

for any n. Then,

$$\begin{aligned} m^*(A) &= m^*\left(A \cap \left(\bigcup_{i=1}^{n} E_i\right)\right) + m^*\left(A \cap \left(\bigcup_{i=1}^{n} E_i\right)^c\right) \quad \text{since } \mathcal{F} \text{ is an algebra} \\ &\geq m^*\left(A \cap \left(\bigcup_{i=1}^{n} E_i\right)\right) + m^*(A \cap E^c) \text{ by the observation above.} \end{aligned}$$

By the Claim above, we have

$$m^*\left(A \cap \left(\bigcup_{i=1}^{n} E_i\right)\right) + m^*(A \cap E^c) = \sum_{i=1}^{n} m^*(A \cap E_i) + m^*(A \cap E^c).$$

Now if we let $n \to \infty$, we have

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

Hence $m^*(A) \geq m^*(A \cap E) + m^*(A \cap E^c)$ and $\bigcup_{i=1}^{\infty} E_i \in \mathcal{F}$. This proves that $\mathcal{F}$ is a $\sigma-$algebra. ∎

Proposition. All intervals are in $\mathcal{F}$ (Hence $\mathcal{B} \subset \mathcal{F}$).

Proof. We will show that $(a, \infty) \in \mathcal{F}$ for $a \in \mathbb{R}$. Let $A \subset \mathbb{R}$, which without loss of generality can be assumed with $m^*(A) < \infty$. Now, from the definition of m^*, we can find a countable family of open intervals $\{E_k\}$ covering A such that for $\varepsilon > 0$

$$m^*(A) \leq \sum_{k=1}^{\infty} \ell(I_k) < m^*(A) + \varepsilon.$$

For any k, let $I_k = I'_k \cup I''_k$ where $I'_k = I_k \cap (a, \infty)$ and $I''_k = I_k \cap (-\infty, a]$. Then

$$m^*(I_k) = \ell(I_k) = \ell(I'_k) + \ell(I''_k) = m^*(I'_k) + m^*(I''_k).$$

Hence

$$m^*(A \cap (a, \infty)) \leq m^*\left(\left(\bigcup_{k=1}^{\infty} I_k\right) \cap (a, \infty)\right) = m^*\left(\bigcup_{k=1}^{\infty} I'_k\right) \leq \sum_{k=1}^{\infty} m^*(I'_k).$$

Similarly we have have

$$m^*(A \cap (-\infty, a]) \leq \sum_{k=1}^{\infty} m^*(I''_k).$$

Then

$$m^*(A \cap (a, \infty)) + m^*(A \cap (-\infty, a]) \leq \sum_{k=1}^{\infty} [m^*(I'_k) + m^*(I''_k)] = \sum_{k=1}^{\infty} m^*(I_k) \leq m^*(A) + \varepsilon.$$

Hence $(a, \infty) \in \mathcal{F}$. Proof for other types of intervals are similar and left as an exercise. ∎

Remark. It follows from this proposition that the σ-algebra $\mathcal{F}$ is rather large (at least as large as $\mathcal{B}$). We will see below that

$$\mathcal{B} \subset \mathcal{F} \subset \mathcal{P}(\mathbb{R}),$$

an each of these inclusions is proper.

Exercise. Let $E \subset \mathbb{R}$ with $m^*(E) > 0$. Show that there exists a bounded subset of E that also has positive outer measure.

Definition. The set function $m = m^*|_{\mathcal{F}} : \mathcal{F} \to [0, \infty]$ is called the **Lebesgue measure** and the triple $(\mathbb{R}, \mathcal{F}, m)$ is sometimes called the **Lebesgue measure space**.

Fact.7 The Lebesgue measure m inherits all of its properties from m^*, namely:

i) $m(\emptyset) = 0$, $m(\mathbb{R}) = \infty$
ii) If $A \subset B$, then $m(A) \leq m(B)$ for all $A, B \in \mathcal{F}$
iii) $m(A \cup B) = m(A) + m(B)$ for any $A, B \in \mathcal{F}$ with $A \cap B = \emptyset$ (finite additivity)
iv) $m(A) = \ell(A)$ if A is an interval
v) m is translation invariant

Furthermore, m is countably additive, namely,

vi) $m(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} m(A_i)$ for any countable disjoint collection $\{A_i\} \subset \mathcal{F}$.

Proof. All these properties, except (iii), (v) and (vi), easily follow from the properties of the Lebesgue outer measure. Since$A \cap B = \emptyset$, it follows that $A^c \cap B = B$. Hence, since A is measurable,

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

which proves (iii). Proof of (v) is straightforward (exercise). For (vi), first observe that from the claim in Theorem.1 it follows that m^* is finitely additive on $\mathcal{F}$, which together with Fact.5 implies (vi). ∎

Another proof goes as follows: We need to show that if $\{E_k\}_{k=1}^{\infty} \subset \mathcal{F}$ is a disjoint collection, then

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

We already know (from finitely additivity) that, for any $n \geq 1$

$$m\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} m(E_k).$$

So by monotonicity

$$m\left(\bigcup_{k=1}^{\infty} E_k\right) \geq m\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} m(E_k).$$

Now if we let $n \to \infty$ we get the desired result. ■

Now, we will exhibit some important features of Lebesgue measurable sets and Lebesgue measure. Let's begin with the approximation property of the Lebesgue measurable sets.

Theorem.2 Let $E \subset \mathbb{R}$. The following are equivalent:

(i) $E \in \mathcal{F}$
(ii) For all $\varepsilon > 0$, there exists O open such that $E \subset O$ and $m^*(O \setminus E) < \varepsilon$.
(iii) For all $\varepsilon > 0$, there exists C closed such that $C \subset E$ and $m^*(E \setminus C) < \varepsilon$.

(iv) There exists $G \in \mathcal{G}_\delta$ with $E \subset G$ such that $m^*(G \setminus E) = 0$.
(v) There exists $F \subset \mathcal{F}_\sigma$ with $F \subset E$ such that $m^*(E \setminus F) = 0$.

Furthermore, if $E \in \mathcal{F}$ with $m(E) < \infty$, then (ii)-(v) are equivalent to

(vi) For all $\varepsilon > 0$, there exists a finite union U of open intervals s.t. $m^*(U \triangle E) < \varepsilon$ where $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

Proof. First we make an important observation: if $m^*(A) < \infty$ is measurable and $A \subset B$, then
$$m^*(B) = m^*(A \cup (B \setminus A)) = m^*(A) + m^*(B \setminus A) \Rightarrow m^*(B \setminus A) = m^*(B) - m^*(A).$$

We will prove (i) $\Rightarrow$ (ii) $\Rightarrow$ (iv) $\Rightarrow$ (i) and (ii) $\Rightarrow$ (iv) $\Rightarrow$ (i). The other implications will be left as an exercise.

(i) $\Rightarrow$ (ii): Assume first that $m^*(E) < \infty$. From definition, given $\varepsilon > 0$ there exists an open cover $\{I_k\}$ of E such that,
$$\sum_{k=1}^{\infty} l(I_k) < m^*(E) + \varepsilon.$$
Let $O = \cup_{k=1}^{\infty} I_k$, then, since m^* is an outer measure,
$$m^*(O) = m^*\left(\bigcup_{k=1}^{\infty} I_k\right) < m^*(E) + \varepsilon.$$
Hence $m^*(O) - m^*(E) < \varepsilon$, and since E has finite outer measure, we have
$$m^*(O \setminus E) = m^*(O) - m^*(E) < \varepsilon.$$

Next let $m^*(E) = \infty$ and let $E_k = E \cap [-k, k]$ for any $k \geq 1$. So $m^*(E_k) < \infty$. Hence by above, there exists O_k open containing E_k such that
$$m^*(O_k \setminus E_k) < \frac{\varepsilon}{2^k}.$$
Since $E = \bigcup_{k=1}^{\infty} E_k \subset \bigcup_{k=1}^{\infty} O_k = O$, where O is open. Then
$$m^*(O \setminus E) == m^*\left(\left(\bigcup_{k=1}^{\infty} O_k\right) \setminus \left(\bigcup_{k=1}^{\infty} E_k\right)\right).$$

Claim. $(\bigcup_{k=1}^{\infty} E_k) \setminus (\bigcup_{k=1}^{\infty} E_k) \subset \bigcup_{k=1}^{\infty}(O_k \setminus E_k)$. (Exercise: Prove the claim.)

Hence
$$m^*(O \setminus E) \leq m^*\left(\bigcup_{k=1}^{\infty}(O_k \setminus E_k)\right) \leq \sum_{k=1}^{\infty} m^*(O_k \setminus E_k) < \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon.$$

(ii) $\Rightarrow$ iv): Given E, by (ii), pick an open set O_k such that $m^*(O_k \setminus E) = \frac{1}{k}$. Now let $G = \bigcap_{k=1}^{\infty} O_k$, then $G \in \mathcal{G}_\delta$. Since $E \subset O_k$ for all k, $E \subset G$ and we have
$$G \setminus E = \left(\bigcap_{k=1}^{\infty} O_k\right) \setminus E = \bigcap_{k=1}^{\infty}(O_k \setminus E) \subset O_k \setminus E$$
for any k. Then $m^*(G \setminus E) \leq m^*(O_k \setminus E) < \frac{1}{k}$ for any $k \geq 1$.. This implies that $m^*(G \setminus E) = 0$.

(iv) $\Rightarrow$ i): For any set $E \subset G$, we have $G = E \cup (G \setminus E)$ and $E = G \cap (G \setminus E)^c$. If G is a set satisfying (iv), then from the fact that $\mathcal{G}_\delta \subset \mathcal{F}$ we have $G \in \mathcal{F}$. Also,
$$m^*(G \setminus E) = 0 \Rightarrow G \setminus E \in \mathcal{F} \Rightarrow (G \setminus E)^c \in \mathcal{F}.$$

Hence $E = G \cap (G \setminus E)^c \in \mathcal{F}$.

(ii) $\Rightarrow$ (vi): Since $E \in \mathcal{F}$, there exists G open with $G \supset E$ such that $m^*(G \setminus E) < \varepsilon/2$. Since G is open, $G = \bigcup_{(k \text{ disjoint})} I_k$. Then

$$\bigcup_{k \text{ disjoint}} I_k = E \cup \left(\bigcup_{k \text{ disjoint}} I_k \setminus E \right) \Rightarrow m^* \left(\bigcup_{k \text{ disjoint}} I_k \right) < m^*(E) + \frac{\varepsilon}{2}.$$

Note that

$$m^* \left(\bigcup_{k \text{ disjoint}} I_k \right) = \sum_k m^*(I_k) = \sum_k \ell(I_k).$$

So $\sum_k \ell(I_k)$ is convergent. Hence there exists N such that

$$\sum_{k=N+1}^{\infty} \ell(I_k) < \frac{\varepsilon}{2}.$$

Now let $U = \bigcup_{k=1}^N I_k$. Then $U \setminus E \subset G \setminus E \Rightarrow m^*(U \setminus E) < \frac{\varepsilon}{2}$. Thus

$$E \setminus U = E \cap U^c = E \cap \left(\bigcup_{k=1}^N I_k \right)^c = E \cap \left(\bigcap_{k=1}^N I_k^c \right).$$

Hence $E \setminus U = \bigcup_{k=1}^N (E \setminus I_k) \subset \bigcup_{k=N+1}^\infty I_k$ and

$$m^*(E \setminus U) < m^* \left(\bigcup_{k=N+1}^{\infty} I_k \right) \leq \sum_{k=N+1}^{\infty} \ell(I_k) < \frac{\varepsilon}{2}.$$

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

(vi) $\Rightarrow$ (i): We begin by making two observations. First, for any sets $U, V \subset \mathbb{R}$, we have $U^c \setminus V^c = V \setminus U$ (and $V^c \setminus U^c = U \setminus V$); hence, $U \vartriangle V = U^c \vartriangle V^c$. Consequently, $m^*(U \vartriangle V) = m^*(U^c \vartriangle V^c)$.

Secondly, if $m^*(U) \geq m^*(V)$, then $U \subset V \cup (U \vartriangle V)$, which implies that $m^*(U) \leq m^*(V) + m^*(U \vartriangle V)$. Similarly, if $m^*(V) \geq m^*(U)$, then $V \subset U \cup (V \vartriangle U)$, and therefore, $m^*(V) \leq m^*(U) + m^*(U \vartriangle V)$. Thus, $|m^*(U) - m^*(V)| \leq m^*(U \vartriangle V)$.

Now, assume (vi); so, given $\epsilon > 0$, let B be an open set (finite union of open intervals) such that $m^*(A \vartriangle B) < \frac{\epsilon}{2}$. Hence, we have $|m^*(A) - m^*(B|) < \frac{\epsilon}{2}$ and $|m^*(A^c) - m^*(B^c)| < \frac{\epsilon}{2}$. Since B is measurable, we have, for any $X \subset \mathbb{R}$, $m^*(X) = m^*(B \cap X) + m^*(B^c \cap X)$. Thus,

$$\begin{aligned}
&|m^*(A \cap X) + m^*(A^c \cap X) - m^*(X)| \\
&= |m^*(A \cap X) + m^*(A^c \cap X) - m^*(B \cap X) + m^*(B^c \cap X)| \\
&\leq |m^*(A \cap X) - m^*(B \cap X)| + |m^*(A^c \cap X) - m^*(B^c \cap X)| \\
&\leq m^*((A \vartriangle B) \cap X) + m^*((A^c \vartriangle B^c) \cap X) \leq m^*(A \vartriangle B) + m^*(A^c \vartriangle B^c) < \epsilon.
\end{aligned}$$

Therefore, $m^*(A \cap X) + m^*(A^c \cap X) = m^*(X)$, implying that $A \in \mathcal{F}$. ∎

Corollary. For all $E \in \mathcal{F}$, $E = B \cup N$ where $B \in \mathcal{B}$ and $m^*(N) = 0$.

Proof. By the previous theorem (v), there exists $F \in F_\sigma$ such that $m^*(E \setminus F) = 0$ and $F \subset E$. Then $E = F \cup (E \setminus F)$. Take $B = F$ and $N = E \setminus F$. ∎

This corollary says that the "completion" of $\mathcal{B}$ is $\mathcal{F}$.

Remark. The analysis above also proves that on (all of) $\mathcal{P}(\mathbb{R})$ Lebesgue outer measure cannot be additive (hence, cannot be countably additive); however, it's additive on a (proper) subset of it, i.e., on $\mathcal{F}$.

Exercises. 1. Prove that $A \in \mathcal{F}$ if and only if for any $\epsilon > 0$ there is a closed set C and an open set O with $C \subset A \subset O$ such that $m^*(O \setminus C) < \epsilon$.

2. Let $E \subset \mathbb{R}$ have finite outer measure. Prove that $E \in \mathcal{F}$ if and only if for each bounded interval (a, b),

$$b - a = m^*((a, b) \cap E) + m^*((a, b) \setminus E).$$

Theorem.3 (Continuity of m, or Monotone convergence theorem for Lebesgue measure) Let $\{E_k\} \subset \mathcal{F}$.

i) If $E_k \subset E_{k+1}$ for all $k \geq 1$, then $m(\bigcup_{k=1}^{\infty} E_k) = \lim_{n\to\infty} m(E_n)$.
ii) If $E_k \supset E_{k+1}$ for all $k \geq 1$ and $m(E_1) < \infty$, then $m(\bigcap_{k=1}^{\infty} E_k) = \lim_{n\to\infty} m(E_n)$.

Proof. (i) Assume $m(E_k) < \infty$ for all $k \geq 1$. Let

$$A_1 = E_1,\ A_2 = E_2 \setminus E_1,\ A_3 = E_3 \setminus E_2, \ldots, A_k = E_k \setminus E_{k-1}$$

for $k \geq 2$. Then $\{A_k\}$ is a disjoint family in $\mathcal{F}$, and furthermore

$$\bigcup_{k=1}^{\infty} E_k = \bigcup_{k=1}^{\infty} A_k.$$

Hence

$$m\left(\bigcup_{k=1}^{\infty} E_k\right) = m\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} m(A_k) = \lim_{n\to\infty} \sum_{k=1}^{n} m(A_k) = \lim_{n\to\infty} \sum_{k=1}^{n} m(E_k \setminus E_{k-1}).$$

We note that the last sum is telescoping. So, it follows that

$$m\left(\bigcup_{k=1}^{\infty} E_k\right) = \lim_{n\to\infty} m(E_n).$$

If there exists an n_0 such that $m(E_{n_0}) = \infty$, then

$$E_{n_0} \subset \bigcup_{n=1}^{\infty} E_n \Rightarrow m\left(\bigcup_{n=1}^{\infty} E_n\right) \geq m(E_{n_0}) = \infty.$$

When $n \geq n_0$, we have $E_{n_0} \subset E_n$, which implies that $\lim_{n\to\infty} m(E_n) = \infty$.

ii) Let $B_k = E_1 \setminus E_k$ for $k \geq 1$. Then, since $\{E_k\}$ is decreasing, $\{B_k\}$ is an increasing collection in $\mathcal{F}$. Also

$$\bigcup_{k=1}^{\infty} B_k = \bigcup_{k=1}^{\infty} (E_1 \setminus E_k) = E_1 \setminus \bigcap_{k=1}^{\infty} E_k.$$

Therefore, since $m(E_1)$ is finite, we now have

$$m\left(\bigcup_{k=1}^{\infty} B_k\right) = m\left(E_1 \setminus \bigcap_{k=1}^{\infty} E_k\right) = m(E_1) - m\left(\bigcap_{k=1}^{\infty} E_k\right).$$

By (i) and the fact that $m(E_k) < \infty$ for each k, we get

$$m\left(\bigcup_{k=1}^{\infty} B_k\right) = \lim_{n\to\infty} m(B_n) = \lim_{n\to\infty} [m(E_1) - m(E_k)] = m(E_1) - \lim_{n\to\infty} m(E_k).$$

Hence $m(\bigcap_{k=1}^{\infty} E_k) = \lim_{n\to\infty} m(E_n)$. ∎

Remark. The condition $m(E_1) < \infty$ in (ii) above cannot be removed. **Exercise:** Give an example that if this condition is removed than the assertion of (ii) is no longer valid.

At this point, we would like to define a set operation that will be studied below in a special setting. Given a collection of subsets $\{E_n\}_n^\infty$ of a set X, similarly to the way we defined the $\limsup a_n$ and $\liminf a_n$ of a sequence, we define

$$\limsup_n A_n = \cap_{n=1}^\infty \cup_{k=n}^\infty A_k, \text{ and}$$
$$\liminf_n A_n = \cup_{n=1}^\infty \cap_{k=n}^\infty A_k.$$

Exercises. Let $\{A_n\}_n \subset \mathcal{F}$.

1. $m(\liminf_n A_n) \leq \liminf_n m(A_n)$.
2. If $m(\cup_{n=1}^\infty A_n) < \infty$, then $m(\limsup_n m(A_n) \leq m(\limsup_n A_n)$.

Definition. We say that a property P holds **almost everywhere** on a measurable set E if there exists $E_0 \subset E$ with $m(E_0) = 0$ such that P holds for all $x \in E \setminus E_0$.

Observation: Let $\{E_k\}_{k=1}^\infty \subset \mathcal{F}$. A point $x \in \mathbb{R}$ belongs to infinitely many $E_k's$ $\Leftrightarrow$ for all $n \geq 1$ there exists $k \geq n$ such that $x \in E_k$ $\Leftrightarrow$ for all $n \geq 1$, $x \in \bigcup_{k\geq n} E_k$ $\Leftrightarrow$

$$x \in \bigcap_{n=1}^\infty \left(\bigcup_{k\geq n} E_k \right) (= \limsup_n E_n).$$

Theorem.4 (Borel-Cantelli Lemma) Let $\{E_k\}_{k=1}^\infty \subseteq \mathcal{F}$ be a countable collection such that

$$\sum_{k=1}^\infty m(E_k) < \infty,$$

then $m(\limsup_n E_n) = 0$; i.e., almost every $x \in \mathbb{R}$ belongs to at most finitely many $E_k's$.

Proof. By the above observation, we need to only show that the set of $x \in \mathbb{R}$ which belong to infinitely many $E_k's$ has measure 0. Equivalently,

$$m\left(\bigcap_{k=1}^\infty \left(\bigcup_{k\geq n} E_k \right) \right) = 0.$$

First, notice that

$$m\left(\bigcup_{k=1}^\infty E_k \right) \leq \sum_{k=1}^\infty m(E_k) < \infty.$$

Lets "decreasify" by letting

$$A_1 = \bigcup_{k=1}^\infty E_k,\ A_2 = \bigcup_{k=2}^\infty E_k,\ \ldots,\ A_n = \bigcup_{k=n}^\infty E_k, \ldots$$

Then $A_k \downarrow$, and $m(A_1) < \infty$. Therefore, by continuity of m,

$$\begin{aligned} m\left(\bigcap_{n=1}^\infty \left(\bigcup_{k\geq n} E_k \right) \right) &= m\left(\bigcap_{n=1}^\infty A_n \right) = \lim_{n\to\infty} m(A_n) \\ &= \lim_{n\to\infty} m\left(\bigcup_{k=n}^\infty E_k \right) \leq \lim_{n\to\infty} \left[\sum_{k=1}^\infty m(E_k) \right] = 0. \quad \blacksquare \end{aligned}$$

Exercises

1. Show that $m^*(A) = m^*(\alpha + A)$ for all $A \subset \mathbb{R},\ \alpha \in \mathbb{R}$.

2. Calculate $m^*(A)$ if

(i) A is a singleton or a finite set,
(ii) $A = \mathbb{Q} \cap [0,1]$.

3. Show that

(i) For any $a, b \in \mathbb{R},\ \ a < b$, the sets $(a,b),\ [a,b),\ (a,b],$ and $[a,b]$ are measurable.
(ii) $\mathbb{Q}$ and $\mathbb{Q}^c$ are measurable, and calculate $m(\mathbb{Q}^c \cap [0,1])$.

4.Let $E \subset \mathbb{R}$ with $m^*(E) > 0$. Show that there exists a bounded subset of E that also has positive outer measure.

5. Show that if $A, B \in \mathcal{F}$, then $m(A \cup B) + m(A \cap B) = m(A) + m(B)$.

6. Prove (i) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (v) in Theorem.2; that is, TFAE:

(i) $E \in \mathcal{F}$
(iii) For all $\varepsilon > 0$, there exists C closed such that $C \subset E$ and $m^*(E \setminus C) < \varepsilon$.
(v) There exists $F \subset \mathcal{F}_\sigma$ with $F \subset E$ such that $m^*(E \setminus F) = 0$.

[Hint: Use the fact that (i) $\Leftrightarrow$ (ii).]

7. Prove that $A \in \mathcal{F}$ if and only if for any $\epsilon > 0$ there is a closed set C and an open set O with $C \subset A \subset O$ such that $m^*(O \setminus C) < \epsilon$.

8. Let $E \subset \mathbb{R}$ have finite outer measure. Prove that $E \in \mathcal{F}$ if and only if for each bounded interval (a,b),

$$b - a = m^*((a,b) \cap E) + m^*((a,b) \setminus E).$$

9. Given a collection of subsets $\{E_n\}_n^\infty$ of a set X, similarly to the way we defined the $\limsup a_n$ and $\liminf a_n$ of a sequence, we define

$$\limsup_n A_n = \cap_{n=1}^\infty \cup_{k=n}^\infty A_k, \text{ and}$$
$$\liminf_n A_n = \cup_{n=1}^\infty \cap_{k=n}^\infty A_k.$$

10. Show that for $\{A_n\}_n \subset \mathcal{F}$.

(i) $m(\liminf_n A_n) \leq \liminf_n m(A_n)$.
(ii) If $m(\cup_{n=1}^\infty A_n) < \infty$, then $m(\limsup_n m(A_n) \leq m(\limsup_n A_n)$.

11. Let $A \in \mathcal{F}$ and $r \in \mathbb{R}$. Show that the sets $r + A$ and rA are also measurable and

$$m(r + A) = m(A), \text{ and } m(rA) = |r| m(A).$$

12. Show that if $A, B \in \mathcal{F}$, then $m(A \cup B) + m(A \cap B) = m(A) + m(B)$.

II.3. Non Measurable Sets

As we promised in section I.1 above, we will now “construct ”a non-measurable set; hence, proving that not all subsets of $\mathbb{R}$ belong to $\mathcal{F}$ (equivalently, not every subset of $\mathbb{R}$ is constructible).

Lemma. Let $E \in \mathcal{F}$ be a bounded set with $m(E) \geq 0$. If Λ is a countably infinite bounded set of real numbers such that $\{\lambda + E\}_{\lambda \in \Lambda}$ is disjoint, then $m(E) = 0$.

Proof. Exercise.

For a set $E \in \mathcal{F}$ let a relation $\sim$ on E be defined by

$$x \sim y \text{ if and only if } x - y \in \mathbb{Q}.$$

It is easy to show that $\sim$ is an equivalence relation on E (Exercise).

Recall that an equivalence relation partitions a set into equivalence classes, the equivalence classes of E with respect to $\sim$ are either disjoint or the same. Namely, $E = \{[x] : [x] \text{ is the equivalence class of } x \text{ with respect to } \sim\}$; and, for all $x, y \in E$, either $[x] \cap [y] = \emptyset$ or $[x] = [y]$. Now, let

$$C_E = \{\text{set of single element representatives from each } E/\sim\}.$$

Then we observe that:

(i) For all $x \in E$, there exists $c \in C_E$ such that $x = c + q$ for some $q \in \mathbb{Q}$.
(ii) For all $x, y \in C_E$, $x \neq y$, $x - y \in \mathbb{Q}^c$.

Hence, for any pair of distinct rationals, say p, q, we have $(p + C_E) \cap (q + C_E) = \emptyset$, for otherwise, it follows that $p + x = q + y$ for some $x, y \in C_E$, which would imply $x - y = p - q \in \mathbb{Q}$, contradicting (ii) above.

Theorem.5 (Vitali) If $E \subset \mathbb{R}$ with $m^*(E) > 0$, then there exists $A \subset E$ such that $A \notin \mathcal{F}$.

Proof. Since $m^*(E) > 0$, we can assume that E is bounded. So consider C_E for this set E.

Claim. $C_E \notin \mathcal{F}$
This claim proves the assertion by letting $A = C_E$. Hence, all we need is to prove the Claim.

For, assume for a contradiction that $C_E \in \mathcal{F}$. Since E is bounded, $E \subset [-m, m]$ for some $m \in \mathbb{Z}$.

Pick $\Lambda_0 = [-2m, 2m] \cap \mathbb{Q}$. Then Λ_0 is countably infinite, bounded, and dense in $[-2m, 2m]$. Then, for any $x \in E$, there exists $c \in C_E$ such that $x = c + q$ for some $q \in \mathbb{Q}$. Since $x, c \in [-m, m]$, $q \in [-2m, 2m]$,

$$\bigcup_{\lambda \in \Lambda_0} (\lambda + C_E)$$

is bounded. By previous Lemma, $m(C_E) = 0$. Hence

$$0 < m^*(E) \leq m^*\left(\bigcup_{\lambda \in \Lambda_0} (\lambda + C_E)\right) \leq \sum_{\lambda \in \Lambda_0} m^*(\lambda + C_E) = \sum_{\lambda \in \Lambda_0} m^*(C_E) = 0.$$

This is a contradiction! Hence the claim must hold. ■

Corollary. There exists $A, B \subset \mathbb{R}$, $A \cap B = \emptyset$, such that

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

Proof. For otherwise Theorem.5 would be false. ■

Remark. It follows from the Corollary that $\mathcal{F} \subsetneq \mathcal{P}(\mathbb{R})$.

Exercise. Let $E \subset \mathbb{R}$ be a non-measurable set with $m^*(E) < \infty$. Show that there is a G_δ-set G with $E \subset G$ such that $m^*(E) = m^*(G)$ and $m^*(G \setminus E) > 0$.

II.4 The Miraculous Cantor Set

First, let's recall the construction of the Cantor set:

$$\begin{aligned} C_0 &= [0,1] \\ C_1 &= \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right] \\ C_2 &= \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right] \\ &\vdots \end{aligned}$$

Let $\mathcal{C} = \bigcap_{n=1}^{\infty} C_n$. Then we have $C_n \supset C_{n+1}$, for all $n \geq 1$. Thus by the Nested Set Theorem, $\mathcal{C} \neq \emptyset$. The set $\mathcal{C}$ is called **the Cantor Set.**

Fact.8 (i) $\mathcal{C}$ is closed.

(ii) $m(\mathcal{C}) = 0$.

(iii) $\mathcal{C}$ is uncountable.

Proof. By construction $\mathcal{C}$ is closed. Next, $\mathcal{C} \in \mathcal{F}$ (why?). Also, since $m(C_1) = \frac{2}{3}$, $m(C_2) = \frac{4}{9}$, $m(C_3) = \frac{8}{27}$, ... , we have $m(C_n) = (\frac{2}{3})^n$ for $n \geq 1$. Hence by the continuity of m,

$$m(\mathcal{C}) = m\left(\bigcap_{n=1}^{\infty} C_n\right) = \lim_{n \to \infty} m(C_n) = 0,$$

proving (i).

Now, observe that if $x \in [0,1]$ can be expressed in ternary form as

$$x = \sum_{k=1}^{\infty} \frac{a_k}{3^k}, \text{ where } a_k \in \{0, 2\},$$

then $x \in \mathcal{C}$. The converse is true as well. Recall that this (ternary) expansion is unique with the convention that, if

$$x = \sum_{k=1}^{\infty} \frac{a_k}{3^k}, \quad \text{where } a_m = 1 \text{ and } a_k = 0 \text{ for } k \geq m+1,$$

then instead take its equivalent

$$x = \sum_{k=1}^{m-1} \frac{a_k}{3^k} + \sum_{k=m}^{\infty} \frac{2}{3^k}.$$

Therefore, it follows that $\mathcal{C} \sim \{0,2\}^{\mathbb{N}}$; and consequently $|\mathcal{C}| = |\{0,2\}^{\mathbb{N}}| = c$. ∎

Exercise. Prove that $\mathcal{C} \oplus \mathcal{C} = [0,2]$ and $\mathcal{C} \ominus \mathcal{C} = [-1,1]$.

Now we will construct a function that will be instrumental in many examples/constructions; namely, the *Cantor-Lebesgue Function.*

Let $\varphi_1 \in (C[0,1], ||\cdot||_\infty)$ be defined as

$$\varphi_1(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x = 1 \\ \frac{1}{2} & \text{if } x \in \left(\frac{1}{3}, \frac{2}{3}\right) \\ \text{interpolate linearly on the rest.} & \end{cases}$$

So φ_1 is continuous, non-decreasing, and $\varphi_1([0,1]) = [0,1]$. Next define

$$\varphi_2(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x = 1 \\ \frac{1}{2} & \text{if } x \in \left(\frac{1}{3}, \frac{2}{3}\right) \\ \frac{1}{4} & \text{if } x \in \left(\frac{1}{9}, \frac{2}{9}\right) \\ \frac{3}{4} & \text{if } x \in \left(\frac{7}{9}, \frac{8}{9}\right) \\ \text{interpolate linearly on the rest.} & \end{cases}$$

Again, φ_2 is continuous, non-decreasing, and $\varphi_2([0,1]) = [0,1]$. Continue constructing $\varphi_n,\ n \geq 3$, in this manner. So $\{\varphi_n\} \subset C[0,1]$ for each $n \geq 1$, defined as

$$\varphi_n(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x = 1 \\ \frac{2k-1}{2^n} & \text{if } \ x \in I_k^n, \quad 1 \leq k \leq 2^{n-1} \\ \text{interpolate linearly on the rest,} & \end{cases}$$

where $\{I_k^n\}_{1\leq k \leq 2^{n-1}}$, are the open intervals removed from the set C_{n-1} to obtain C_n in the construction of Cantor set. Now, for $m < n$ are integers, we have

$$|\varphi_n(x) - \varphi_m(x)| < \frac{1}{2^m}$$

for all $x \in [0,1]$ (notice that $|\varphi_{n+1} - \varphi_n| \leq \frac{1}{2}|\varphi_n - \varphi_{n-1}|$ for $n \geq 2$). Hence

$$||\varphi_n - \varphi_m||_\infty = \max_{x\in[0,1]} |\varphi_n(x) - \varphi_m(x)| < \frac{1}{2^m};$$

and hence, $\{\varphi_m\}$ is Cauchy in $C[0,1]$. Since $C[0,1]$ is a complete metric space, $\{\varphi_n\}$ converges uniformly to a function φ, that is continuous and non-decreasing with $\varphi([0,1]) = [0,1]$. Thus $\varphi : [0,1] \to [0,1]$ is an onto function such that

- φ is non-decreasing,
- $\varphi(x) = \frac{k}{2^n}$, $n \geq 1$. $1 \leq k \leq 2^{n-1}$ when $x \in I_n^k$, where I_n^k's are intervals removed at the n^{th} stage of the Cantor set construction.

Let $O = [0,1] \setminus \mathcal{C}$, which is measurable. Then $[0,1] = O \cup \mathcal{C}$ and

$$[0,1] = \varphi([0,1]) = \varphi(\mathcal{C}) \cup \varphi(O).$$

Since $\varphi(O)$ is a countable set, $m^*(\varphi(O)) = 0$. Hence, $\varphi(O) \in \mathcal{F}$. Consequently, $\varphi(\mathcal{C})$ is also measurable with measure $m(\varphi(\mathcal{C})) = 1$. Thus by Vitali's Theorem, there exists $P \subset \varphi(\mathcal{C})$ such that $P \notin \mathcal{F}$. On the other hand $\varphi^{-1}(P) \subset \mathcal{C}$. Since $m(\varphi^{-1}(P)) = 0$, it follows that $\varphi^{-1}(P) \in \mathcal{F}$. Thus, we proved that

Fact.9 The function φ maps a measurable set (a subset of $\mathcal{C}$) onto a non-measurable set.

Next let $f : [0,1] \to [0,2]$ by $f(x) = x + \varphi(x)$. Then it follows from the properties of φ and the function x that

- f is continuous and onto $[0,2]$.
- f is strictly increasing.

Since f is strictly increasing, it maps Borel sets onto Borel sets. (Why?) Note that if $A \in \mathcal{F}$ such that $\varphi(A) \notin \mathcal{F}$, then $f(A) \notin \mathcal{F}$ either. Then $A \notin \mathcal{B}$ holds, for otherwise if $A \in \mathcal{B}$, so would $f(A)$. This implies that $f(A) \in \mathcal{F}$ which is a contradiction. Hence:

Corollary. There exists a nonempty set $A \in \mathcal{F} \setminus \mathcal{B}$.

Exercise. Show that there exists a continuous, strictly increasing function $f : [0,1] \to \mathbb{R}$ that maps a set of positive measure onto a set of measure zero.

II.5. General Measure Spaces (A Brief Survey)

In this section, following similar steps as in Section II.2, we will construct measures on sets not necessarily equal to $\mathbb{R}$, generalizing the Lebesgue measure.

Definition. A nonempty collection $\mathcal{G}$ of subsets of a set X is called an **algebra** if

i) $X \in \mathcal{G}$,
ii) if $\{A_i\}_{i=1}^n \subset \mathcal{G}$, then $\cup_{i=1}^n A_i \in \mathcal{G}$, and
iii) $A \in \mathcal{G}$ implies $X \setminus A \in \mathcal{G}$.

Remark. It follows that $\emptyset \in \mathcal{G}$ and that if $\{A_i\}_{i=1}^n \subset \mathcal{G}$, then $\cap_{i=1}^n A_i \in \mathcal{G}$.

Definition. A nonempty collection $\mathcal{G}$ of subsets of X is called a **σ-algebra** if it is an algebra and (ii′) if $\{A_i\}_{i=1}^\infty \subset \mathcal{G}$, then $\cup_{i=1}^\infty A_i \in \mathcal{G}$.

Remark. The collection $\mathcal{G} = \{\emptyset,\ X\}$ and $\mathcal{P}(X)$ are σ-algebras for any X, where $\{\emptyset,\ X\}$ is the smallest and $\mathcal{P}(X)$ is the largest σ-algebra.

It is easy to see that the intersection of a family $\{\mathcal{A}_\alpha\}$ of $\sigma-$algebras of a set X is another $\sigma-$algebra (Exercise). As before, we actually have a more general statement:

Fact.10 Given a collection $\mathcal{E}$ of subsets of X, then there exists a smallest $\sigma-$algebra $\mathcal{A}$ of subsets of X such that $\mathcal{E} \subset \mathcal{A}$.

Proof. Exercise. [Hint: Mimic the proof of Fact.1 in Section II.1]

Notation: $\mathcal{A} := \sigma(\mathcal{E})$ is called as the **$\sigma-$algebra generated by $\mathcal{E}$**.

Exercise. If $\mathcal{G}$ is a $\sigma-$algebra and $E \in \mathcal{G}$, then the collection $\mathcal{A} = \{A \cap E : A \in \mathcal{G}\}$ is also a $\sigma-$algebra.

Definition. Let X be a (non-empty) set and $\mathcal{A}$ be a $\sigma-$algebra of subsets of X, then the pair $(X, \mathcal{A})$ is called a **measurable space**.

Examples. 1. Let X be any uncountable set and let

$$\mathcal{A} = \{A \subset X : |A| \leq \aleph_0 \text{ or } |A^c| \leq \aleph_0\}.$$

Then $\mathcal{A}$ is a σ-algebra and $(X, \mathcal{A})$ is a measurable space.

2. Let $X = \mathbb{R}$ and

$$\mathcal{A} = \{\text{finite, disjoint union of sets } [a,b),\ \text{where } -\infty \leq a < \infty,\ -\infty < b \leq \infty\}.$$

Then $(X, \mathcal{A})$ is not a measurable space. Note that $\mathcal{A}$ is an algebra, but not a $\sigma-$algebra. On the other hand, $\sigma(\mathcal{A}) = \mathcal{B}$, which is a σ-algebra.

Definition. Let $(X, \mathcal{A})$ be a measurable space. A set function $\mu : \mathcal{A} \to [0, \infty]$ is called a **measure** (on $\mathcal{A}$) if:

i) $\mu(\emptyset) = 0$.
ii) $\mu(\bigcup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} \mu(E_k)$ for any disjoint collection $\{E_k\} \subset \mathcal{A}$.

The triple $(X, \mathcal{A}, \mu)$ is called a **measure space**. If $\mu(X) < \infty$, then $(X, \mathcal{A}, \mu)$ is called a **finite measure space**; if, in particular, $\mu(X) = 1$, then $(X, \mathcal{A}, \mu)$ is called a **probability space**. If $X = \bigcup_{n=1}^{\infty} X_n$ such that $\mu(X_n) < \infty$ for each n, then $(X, \mathcal{A}, \mu)$ is called a $\boldsymbol{\sigma}$**−finite** measure space.

Definition. A set function $\mu : \mathcal{A} \to [0, \infty]$ satisfying

i) $\mu(\emptyset) = 0$.
ii) $\mu(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} \mu(E_i)$

for any disjoint collection $\{E_i\}_{i=1}^{n} \subset \mathcal{A}$, is called a **finitely additive measure**.

Examples. 1. $(\mathbb{R}, \mathcal{F}, m)$ and $(\mathbb{R}, \mathcal{B}, m)$ are both measure spaces.

2. For any $a, b \in \mathbb{R},\ a < b$, the triple $([a, b], \mathcal{F}|_{[a,b]}, m)$, where $\mathcal{F}|_{[a,b]} = \{A \cap [a, b] : A \in \mathcal{F}\}$, is a finite measure space; in particular, $([0, 1], \mathcal{F}|_{[0,1]}, m)$ is a probability space.

3. Let X be any infinite set and let $\mathcal{A} = \mathcal{P}(X)$. Define $\mu : \mathcal{A} \to [0, \infty]$ by

$$\mu(A) = \begin{cases} 0 & \text{if } |A| < \infty \\ \infty & \text{if } A \text{ is infinite.} \end{cases}$$

Then μ is <u>not</u> a measure, but is a finitely additive measure.

4. Let $X = \mathbb{Z}$, $\mathcal{A} = \mathcal{P}(\mathbb{Z})$, and define $c : \mathcal{A} \to [0, \infty]$ by $c(A) = |A|$. Then c is a measure on $\mathbb{Z}$ (called the counting measure).

5. Let X be any uncountable set and let

$$\mathcal{A} = \{A \subset X : A \text{ or } A^c \text{ is countable}\}.$$

If a set function $\mu : \mathcal{A} \to [0, \infty]$ is defined by

$$\mu(A) = \begin{cases} 0 & \text{if } A \text{ is countable} \\ 1 & \text{if } A^c \text{ is countable} \end{cases},$$

then μ is a measure.

Remark. Sets of measure zero and almost all properties are defined in any measure space exactly the same as in the case of the Lebesgue measure.

Theorem.6 Let $(X, \mathcal{A}, \mu)$ be a measure space. Then

i) For all $A, B \in \mathcal{A}$ such that $A \subset B$, $\mu(A) \le \mu(B)$.
ii) If $\{A_i\}_{i=1}^{\infty} \subset \mathcal{A}$, then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \le \sum_{i=1}^{\infty} \mu(A_i) \text{ (Subadditivity).}$$

iii) If $\{A_i\} \subset \mathcal{A}$, and $A_i \subset A_{i+1}$ for all $i \ge 1$, then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} \mu(A_n) \text{ (Continuity from below).}$$

iv) If $\{E_i\} \subset \mathcal{A}$, and $E_i \supset E_{i+1}$, for all $i \geq 1$ and $\mu(E_1) < \infty$, then

$$\mu\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} \mu(E_n) \text{ (Continuity from above).}$$

v) Borel-Cantelli Lemma holds.

Proof. Similar to the proofs of the analogous properties in the case of m. ■

Discussion: $(\mathbb{R}, \mathcal{B}, m)$ vs. $(\mathbb{R}, \mathcal{F}, m)$.
It is known that $C \in \mathcal{B}$, but C contains an $E \subset C$ such that E is not a Borel set. Thus, in $(\mathbb{R}, \mathcal{B}, m)$, C has a subset, say E, such that $E \notin \mathcal{B}$. However, $E \in \mathcal{F}$. Notice $m(C) = 0$, hence there exists a measure space with the property that not every subset of measure zero is measurable. Such measure spaces are called *incomplete.*

Definition. A measure space $(X, \mathcal{A}, \mu)$ is called **complete** if $E \in \mathcal{A}$ where $\mu(E) = 0$ and $F \subset E$, then $F \in \mathcal{A}$.

Theorem.7 Let $(X, \mathcal{A}, \mu)$ be a measure space and $\mathcal{N} = \{N \in \mathcal{A} : \mu(N) = 0\}$. Let

$$\overline{\mathcal{A}} = \{E \cup F : E \in \mathcal{A}, F \subset N \text{ for some } N \in \mathcal{N}\},$$

and define $\overline{\mu} : \overline{\mathcal{A}} \to [0, \infty]$ by $\overline{\mu}(E \cup F) = \mu(E)$. Then $\overline{\mathcal{A}}$ is a $\sigma-$algebra and $\overline{\mu}$ is a measure on $\overline{\mathcal{A}}$ such that $(X, \overline{\mathcal{A}}, \overline{\mu})$ is a complete measure space.

Proof. First we will show that $\overline{\mathcal{A}}$ is a $\sigma-$algebra. Let $E = A \cup F \in \overline{\mathcal{A}}$. Without loss of generality we can assume that $A \cap F = \emptyset$ (otherwise replace F by $F \setminus A$ and N by $N \setminus A$).

Observe that $A \cup F = (A \cup N) \cap (N^c \cup F)$. Then

$$E^c = (A \cup N)^c \cup (N^c \cup F)^c = (\underbrace{A \cup N}_{\in \mathcal{A}})^c \cup (\underbrace{N \setminus F}_{\in \mathcal{N}}).$$

Thus $E^c \in \overline{\mathcal{A}}$. If $\{A_i \cup F_i\} \subset \overline{\mathcal{A}}$, then

$$\bigcup_{i=1}^{\infty}(A_i \cup F_i) = (\underbrace{\bigcup_{i=1}^{\infty} A_i}_{\in \mathcal{A}}) \cup (\underbrace{\bigcup_{i=1}^{\infty} F_i}_{\in \mathcal{N}}).$$

Thus $\overline{\mathcal{A}}$ is a $\sigma-$algebra. It is left as an exercise to show that $\overline{\mu}$ satisfies all the properties of being a measure.

Now, all we need to show is that $\overline{\mu}$ is well defined. Let $E \subset \overline{\mathcal{A}}$, and $A_1 \cup F_1 = E = A_2 \cup F_2$. Then $A_1 \subset A_2 \cup F_2 \subset (A_2 \cup N_2)$. Hence

$$\mu(A_1) \leq \mu(A_2 \cup N_2) = \mu(A_2).$$

Similarly we get $\mu(A_1) \geq \mu(A_2)$. Thus $\mu(A_1) = \mu(A_2)$, so $\overline{\mu}$ is well defined. ■

Question: How can we construct measures?

We will answer this question, as in the case of Lebesgue measure, via "outer measures."

Definition. Let $X \neq \emptyset$ be a set. A set function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ is called an **outer measure** (on X) if :

i) $\mu^*(\emptyset) = 0$.
ii) $\mu^*(A) \leq \mu^*(B)$ when $A \subset B$.
iii) $\mu^*(\bigcup_{k=1}^{\infty} A_k) \leq \sum_{k=1}^{\infty} \mu^*(A_k)$.

Exercise. If an outer measure is is finitely additive, then it is a measure.

Given an outer measure on X, then we define μ^*-measurable sets as:
Definition. A set $E \subset X$ is called a μ^*-**measurable set** if for all $A \subset X$,

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

Examples. 1. Let X be an infinite set and define $\mu^*(E) = |E|$ if E is finite and $\mu^*(E) = \infty$ if E is infinite. Then μ^* is an outer measure (Show). Exercise: Determine the μ^*-measurable sets.

2. Let X be any set and define $\mu^*(\emptyset) = 0$ and $\mu^*(E) = 1$ if $E \neq \emptyset$. Then μ^* is an outer measure (Show). Exercise: Determine the μ^*-measurable sets.

3. 1. Let X be an uncountable set and define $\mu^*(E) = 0$ if E is countable and $\mu^*(E) = 1$ if E is uncountable. Then μ^* is an outer measure (Show). Exercise: Determine the μ^*-measurable sets.

Theorem.8 (Carathèodory) Given a set X and an outer measure μ^* on X, the collection

$$\mathcal{A} := \{E \subset X : \ E \ \text{ is } \mu^* - \text{measurable}\}$$

is a $\sigma-$algebra of subsets of X, and $\mu^*|_{\mathcal{A}} = \mu$ is a (complete) measure on X.

Proof. Similar to the corresponding proofs for m (Theorem.1, Theorem.2 and its Corollary in Section II.2). ■

Question: How can we construct outer measures?

Definition. Let $X \neq \emptyset$ be a set. A collection κ of subset of X is called a **(sequential) covering class** for X if for all $A \subset X$, there exists $\{E_i\}_{i=1}^{\infty} \subset \kappa$ such that $A \subset \bigcup_{i=1}^{\infty} E_i$.

Examples.

1. In $\mathbb{R}$, $\kappa = \{\emptyset\} \cup \{(n, n+2)\}_{n \in \mathbb{Z}}$ is a covering class.
2. In $\mathbb{R}$, $\kappa = \{\text{All intervals with rational endpoints}\} \cup \{\emptyset\}$ is a covering class (for m).

.

Exercise. Find a suitable sequential covering class κ for X in each of the examples 1-3 above.

Let $\lambda : \kappa \to [0, \infty]$ be a set function such that $\lambda(\emptyset) = 0$. For any $A \subset X$, let $\mu_\lambda^*(A)$ be the set function (induced by λ) on $\mathcal{P}(X)$ defined as

$$\mu_\lambda^*(A) = \inf \left\{ \sum_{k=1}^{\infty} \lambda(E_k) : \{E_k\} \in \kappa, \bigcup_{k=1}^{\infty} E_k \supset A \right\}.$$

Theorem.9 The set function μ_λ^* defined above is an outer measure on X.

Proof (sketch). $\mu_\lambda^*(\emptyset) = 0$, $\mu_\lambda^*(A) \geq 0$ for all $A \in X$, and $A \subset B \Rightarrow \mu_\lambda^*(A) \leq \mu_\lambda^*(B)$ is left as an exercise to the reader. We need to prove the countable subadditivity property. Let

$\{A_k\} \subset \mathcal{P}(X)$. Given $\epsilon > 0$,

$$\begin{aligned} A_1 &\subset \bigcup_{k=1}^{\infty} E_k^1 \text{ such that } \{E_k^1\} \subset \kappa \text{ with } \sum_{k=1}^{\infty} \lambda(E_k^1) < \mu^*(A_1) + \frac{\epsilon}{2} \\ A_2 &\subset \bigcup_{k=1}^{\infty} E_k^2 \text{ such that } \{E_k^2\} \subset \kappa \text{ with } \sum_{k=1}^{\infty} \lambda(E_k^2) < \mu^*(A_2) + \frac{\epsilon}{2^2} \\ &\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots \end{aligned}$$

Now show that $\mu_\lambda^*(\bigcup_{m=1}^{\infty} A_m) \leq \sum_{m=1}^{\infty} \mu_\lambda^*(A_m) + \epsilon$. Since $\epsilon > 0$ is arbitrary, we get the desired result. ■

Examples. 1. If $X = \mathbb{R}$, $\kappa = \{\text{all open intervals}\}$, and $\lambda = \text{length}$, then we obtain $\mu_\lambda^* = m^*$ which gives the Lebesgue measure m.

2. (Lebesgue Stieltjes measure) Let $X = \mathbb{R}$, $\kappa = \{(-\infty, a] : a \in \mathbb{R}\}$ or $\kappa = \{(a, b] : a, b \in \mathbb{R}^{\#}\}$. Let $f : \mathbb{R} \to \mathbb{R}$ be such that f is monotone increasing and right continuous (i.e. $\lim_{x \to a^+} f(x) = f(a)$). Define $\lambda(a, b] = f(b) - f(a)$ with the convention that $\lambda(\emptyset) = 0$. Let μ_f^* denote the set function "induced by λ" (via f). Then by Theorem.9, μ_f^* is the outer measure induced by f, and by Theorem.8, $\mu_f = \mu_f^*|_{\mathcal{A}=\mathcal{F}} =$ is the induced measure, called the **Lebesgue-Stieltjes measure** induced by f.

Exercises

1. If $\mu_1, \mu_2, \ldots, \mu_n$ are measures on a measurable space $(X, \mathcal{A})$ and $\alpha_1, \alpha_2, \ldots, \alpha_n \in [0, \infty)$, then $\mu = \sum_{i=1}^{n} \alpha_i \mu_i$ is also a measure on $(X, \mathcal{A})$.

2. If $(X, \mathcal{A}, \mu)$ is a measure space and $E \in \mathcal{A}$, then the set function $\mu_E : \mathcal{A} \to [0, \infty]$, defined by $\mu_E(A) = \mu(E \cap A)$ for all $A \in \mathcal{A}$, is also a measure on $\mathcal{A}$.

3. Let μ be a finitely additive measure on a σ-algebra $\mathcal{A}$ of subsets of a set X. μ is a measure if and only if it is continuous from below.

4. Let μ be a finitely additive measure on a σ-algebra $\mathcal{A}$ of subsets of a set X with $\mu(X) < \infty$. μ is a measure if and only if it is continuous from above.

5. Let $(X, \mathcal{A}, \mu)$ be a measure space.

a) If $A, B \in \mathcal{A}$ and $\mu(A \triangle B) = 0$, then $\mu(A) = \mu(B)$.
b) Define: $A \sim B$ if and only if $\mu(A \triangle B) = 0$. Then "$\sim$" is an equivalence relation on $\mathcal{A}$.
c) For $A, B \in \mathcal{A}$, define $\rho(A, B) = \mu(A \triangle B)$. Then ρ is a metric on $\mathcal{A}/\sim$.
(Hence, $(\mathcal{A}/\sim\ ,\ \rho)$ is a metric space).

6. If an outer measure μ^* on a set $X \neq \emptyset$ is finitely additive, then it is a measure.

7. Let $X \neq \emptyset$ and $\kappa = \{\emptyset,\ X,\ \text{one-point sets}\}$. Define $\lambda(X) = \infty$, $\lambda(\emptyset) = 0$, and $\lambda(E) = 1$ if $\emptyset \neq E \subsetneq X$. Describe the outer measure μ_λ^*.

8. Let X be an uncountable set and $\kappa = \{\emptyset,\ X,\ \text{one-point sets}\}$. Define $\lambda(X) = 1$ and $\lambda(E) = 0$ if $E \subsetneq X$. Describe the outer measure μ_λ^*.

9. Let $X \neq \emptyset$, κ be a covering class for X and λ be a set function on $\mathcal{P}(X)$. If $E \in \kappa$, then $\mu_\lambda^*(E) \leq \lambda(E)$. Give an example where strict inequality occurs.

10. Let $X \neq \emptyset$, κ be a covering class for X, λ be a set function on $\mathcal{P}(X)$ and μ_λ^* be the outer measure induced by κ and λ.

a) If κ is a σ-algebra and λ is a measure, then $\mu_\lambda^*(E) = \lambda(E)$ for all $E \in \kappa$.
b) If κ is a σ-algebra and λ is a measure, then every set in κ is μ_λ^*-measurable.

[The measure μ_λ given by μ_λ^* is called an **extension** of λ.]

11. Let $X \neq \emptyset$, $\kappa = \{\emptyset, X\}$ and λ be a set function on $\mathcal{P}(X)$ given by $\lambda(E) = 0$ if $E = \emptyset$ and $\lambda(E) = 0$ if $E = X$. Determine the outer measure μ_λ^* and describe the σ-algebra of μ_λ^*-measurable sets. The same question when $X = [0, 1]$.

12. Let $X = \mathbb{R}$, $\kappa = \mathcal{P}(\mathbb{R})$ and define $\lambda(E) = |\{k \in \mathbb{Z} : \ k \in E\}|$. Determine the outer measure μ_λ^* and describe the σ-algebra of μ_λ^*-measurable sets.

13. Let F be a real-valued, increasing and right continuous function on $\mathbb{R}$. If μ_F is the associated Lebesgue-Stieltjes measure, then prove that, for any $a, b \in \mathbb{R}$,

(i) $\mu_F(\{a\}) = F(a) - F(a_-)$,
(ii) $\mu_F([a, b)) = F(b_-) - F(a_-)$,
(iii) $\mu_F([a, b]) = F(b) - F(a_-)$, and
(iv) $\mu_F((a, b)) = F(b_-) - F(a)$.