Another method of Integration: Lebesgue Integral

Shengjun Wang

2017/05

Abstract

Centuries ago, a French mathematician Henri Lebesgue noticed that the Riemann Integral does not work well on unbounded functions. It leads him to think of another approach to do the integration, which is called Lebesgue Integral. This paper will briefly talk about the inadequacy of the Riemann integral, and introduce a more comprehensive definition of integration, the Lebesgue integral. There are also some discussion on Lebesgue measure, which establish the Lebesgue integral. Some examples, like F_σ set, G_δ set and Cantor function, will also be mentioned.

Contents

1 Riemann Integral **4**

1.1 Riemann Integral 4

1.2 Inadequacies of Riemann Integral 7

2 Lebesgue Measure **8**

2.1 Outer Measure 9

2.2 Measure 12

2.3 Inner Measure 26

3 Lebesgue Integral **28**

3.1 Lebesgue Integral 28

3.2 Riemann vs. Lebesgue 30

3.3 The Measurable Function 32

1 Riemann Integral

1.1 Riemann Integral

In a first year "real analysis" course, we learn about continuous functions and their integrals. Mostly, we will use the Riemann integral, but it does not always work to find the area under a function.

The Riemann integral that we studied in calculus is named after the German mathematician Bernhard Riemann and it is applied to many scientific areas, such as physics and geometry. Since Riemann's time, other kinds of integrals have been defined and studied; however, they are all generalizations of the Riemann integral, and it is hardly possible to understand them or appreciate the reasons for developing them without a thorough understanding of the Riemann integral [1].

Let us recall the definition of the Riemann integral.

Definition 1.1. A **partition** P of an interval $[a, b]$ is a finite set of points $\{x_i : 0 \leq i \leq n\}$ such that

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

Definition 1.2. Let f be a bounded real-valued function defined on the interval $[a, b]$ and let

$$a = x_0 < x_1 < ... < x_n = b$$

be a subdivision of $[a, b]$. For each subdivision we can define the **upper sum** of f over the this subdivision as

$$S = \sum_{i=1}^{n} (x_i - x_{i-1}) M_i$$

and the **lower sum** of f over this subdivision as

$$s = \sum_{i=1}^{n} (x_i - x_{i-1}) m_i,$$

where

$$M_i = \sup_{x_{i-1} \leq x \leq x_i} f(x) \quad \text{and} \quad m_i = \inf_{x_{i-1} \leq x \leq x_i} f(x).$$

Definition 1.3. A function $f : [a, b] \to \mathbb{R}$ is **Riemann integral** on $[a, b]$ if there exists a number L with the following property: for each $\epsilon > 0$ there exists $\delta > 0$ such that

$$|\sigma - L| < \epsilon$$

if σ is any Riemann sum of f over a partition P of $[a, b]$ such that $||P|| < \delta$. In this case, we say that L is the Riemann integral of f over $[a, b]$, and write

$$\int\limits_a^b f(x)dx = L.$$

Then, we define the upper Riemann integral and lower Riemann integral in the following way.

Definition 1.4. The **upper Riemann integral** of f on $[a, b]$ is denoted by

$$(R)\overline{\int_a^b} f(x)\,dx = \inf S$$

and the **lower Riemann integral** of f on $[a, b]$ is denoted by by

$$(R)\underline{\int_a^b} f(x)\,dx = \sup s.$$

Note that the upper Riemann integral of f is always greater than or equal to the lower Riemann integral. When the two are equal to each other, we say that f is Riemann integrable on $[a, b]$, and we call this common value the Riemann integral of f. We denote it by

$$(R)\int\limits_a^b f(x)\,dx,$$

to distinguish it from the Lebesgue integral, which we will encounter later.

We define a step function as a function ψ which has the form

$$\psi(x) = c_i, \quad \text{where} \quad x_{i-1} < x < x_i$$

for some subdivision of $[a, b]$ and some set of constants c_i. We know that this step function is integrable, then

$$(R)\int\limits_a^b \psi(x)\,dx = \sum_{i=1}^n c_i(x_i - x_{i-1}) \quad \text{and} \quad (R)\overline{\int_a^b} \psi(x)\,dx = (R)\underline{\int_a^b} \psi(x)\,dx.$$

With this idea in mind, let us consider the following example.

Example 1. *We will compute the upper Riemann integral and the lower Riemann integral of the following function.*

$$f(x) = \begin{cases} 0, & x \text{ is rational.} \\ 1, & x \text{ is irrational.} \end{cases}$$

And its graph looks like this[2]:

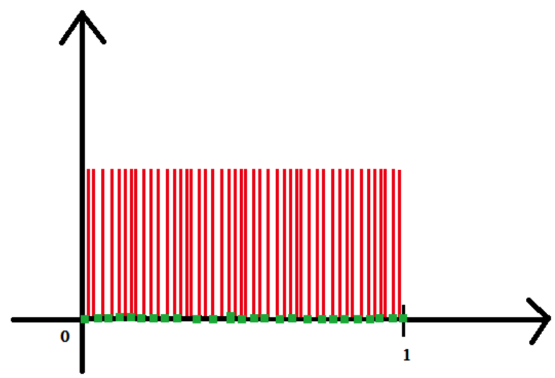

Figure 1: $f(x)$ in Example 1

If we partition the domain of this function, then each subinterval will contain both rational and irrational, since both of them are dense in real. Thus, the supremum on each subinterval is 1 and the infimum on each subinterval is 0. From definition on the previous page, we know that

$$(R)\overline{\int_a^b} f(x)\,dx = b - a \quad \text{and} \quad (R)\underline{\int_a^b} f(x)\,dx = 0.$$

In this case, we realize that this function is not Riemann integrable, because the upper Riemann integral is not equal to the lower Riemann integral. This illustrates one of the shortcomings of the Riemann integral. For some discontinuous functions, we are not able to integrate them with the Riemann integral and measure the areas under those functions.

1.2 Inadequacies of Riemann Integral

From the previous problem, we realized one of the shortcomings of the Riemann integral. For these reasons, we should find another type of integral, which not only corresponds to the Riemann integral, but also covers the non-Riemann integrable functions.

The Riemann integral is based on the fact that by partitioning the domain of an assigned function, we approximate the assigned function by piecewise constant functions in each sub-interval. In contrast, the Lebesgue integral partitions the range of that function.

A great analogy to Lebesgue integration is given in [3]: Suppose we want both student R (Riemann's method) and student L(Lebesgue's method) to give the total value of a bunch of coins with different face values lying on a table. Student R will add up the face value of each coin and come up with the total value as he randomly picks up a coin. In contrast, student L will categorize all the coins with respect to their face value and count the number of coins corresponding to each face value. Multiplying the quantity with its corresponding face value and adding them up, student L will give us the total value.

In calculus, we learned that integration will help us to calculate the length, area, and volume in different dimensions for a given subset of the domain. Now we want to know what is the size of preimage of subset in that range.

2 Lebesgue Measure

Let us consider a bounded interval I with endpoints a and b ($a < b$). The length of this bounded interval I is defined by $\ell(I) = b - a$. In contrast, the length of an unbounded interval, such as (a, ∞), $(-\infty, b)$ or $(-\infty, \infty)$, is defined to be infinite. Obviously, the length of a line segment is easy to quantify. However, what should we do if we want to measure an arbitrary subset of $\mathbb{R}$? The Lebesgue measure, named after Henri Lebesgue, is one of the approaches that helps us to investigate this problem. Given a set E of real numbers, we denote the Lebesgue measure of set E by $\mu(E)$. To correspond with the length of a line segment, the measure of a set A should keep the following properties:

(1) If A is an interval, then $\mu(A) = \ell(A)$.

(2) If $A \subseteq B$, then $\mu(A) \leq \mu(B)$.

(3) Given $A \subseteq \mathbb{R}$ and $x_0 \in \mathbb{R}$, define $A + x_0 = \{x + x_0 : x \in A\}$. Then $\mu(A) = \mu(A + x_0)$.

(4) If A and B are disjoint sets, then $\mu(A \bigcup B) = \mu(A) + \mu(B)$. If $\{A_i\}$ is a sequence of disjoint sets, then $\mu(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} \mu(A_i)$.

However, we might encounter the situation that not all the properties are satisfied. For example, as we will see there exists non-measurable sets. Then we cannot do the measure simply by going on the set operation. Let us put this bizarre situation aside for now and discuss it later. Before we step into the Lebesgue measure, let us define Lebesgue outer measure first.

2.1 Outer Measure

Definition 2.1. Let E be a subset of $\mathbb{R}$. Let $\{I_k\}$ be a sequence of open intervals. The **Lebesgue outer measure** of E is defined by

$$\mu^*(E) = \inf \left\{ \sum_{k=1}^{\infty} \ell(I_k) : E \subseteq \bigcup_{k=1}^{\infty} I_k \right\}.$$

Note that $0 \leq \mu^*(E) \leq \infty$.

Theorem 2.1. *Lebesgue outer measure has the following properties:*

(a) If $E_1 \subseteq E_2$, then $\mu^*(E_1) \leq \mu^*(E_2)$.

(b) The Lebesgue outer measure of any countable set is zero.

(c) The Lebesgue outer measure of the empty set is zero.

(d) Lebesgue outer measure is invariant under translation, that is,

$$\mu^*(E + x_0) = \mu^*(E).$$

(e) Lebesgue outer measure is countably sub-additive, that is,

$$\mu^*(\bigcup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} \mu^*(E_i).$$

(f) For any interval I, $\mu^*(I) = \ell(I)$.

Proof. Part (a) is trivial.

For part (b) and (c), let $E = \{x_k : k \in Z^+\}$ be a countably infinite set. Let $\epsilon > 0$ and let ϵ_k be a sequence of positive numbers such that $\sum\limits_{k=1}^{\infty} \epsilon_k = \frac{\epsilon}{2}$. Since

$$E \subseteq \bigcup_{k=1}^{\infty} (x_k - \epsilon_k, x_k + \epsilon_k),$$

it follows that $\mu^*(E) \leq \epsilon$. Hence, $\mu^*(E) = 0$. Since $\emptyset \subseteq E$, then $\mu^*(\emptyset) = 0$

For part (d), since each cover of E by open intervals can generate a cover of $E + x_0$ by open intervals with the same length, then $\mu^*(E + x_0) \leq \mu^*(E)$. Similarly, $\mu^*(E + x_0) \geq \mu^*(E)$, since E is a translation of $E + x_0$ Therefore, $\mu^*(E + x_0) = \mu^*(E)$.

For part (e), if $\sum\limits_{i=1}^{\infty}\mu^*(E_i)=\infty$, then the statement is trivial. Suppose that the sum is finite and let $\epsilon>0$. For each i, there exists a sequence $\{I_k^i\}$ of open intervals such that $E_i\subseteq\bigcup\limits_{k=1}^{\infty}I_k^i$ and $\sum\limits_{k=1}^{\infty}\ell(I_k^i)<\mu^*(E_i)+\frac{\epsilon}{2^i}$. Now $\{I_k^i\}$ is a double-indexed sequence of open intervals such that $\bigcup\limits_{i=1}^{\infty}E_i\subseteq\bigcup\limits_{i=1}^{\infty}\bigcup\limits_{k=1}^{\infty}I_k^i$ and

$$\sum_{i=1}^{\infty}\sum_{k=1}^{\infty}\ell(I_k^i)<\sum_{i=1}^{\infty}(\mu^*(E_i)+\frac{\epsilon}{2^i})=\sum_{i=1}^{\infty}\mu^*(E_i)+\epsilon.$$

Therefore, $\mu^*(\bigcup\limits_{i=1}^{\infty}E_i)\leq\sum\limits_{i=1}^{\infty}\mu^*(E_i)+\epsilon$. The result follows since $\epsilon>0$ was arbitrary.

For part (f), we need to prove $\mu^*(I)\leq\ell(I)$ and $\mu^*(I)\geq\ell(I)$ respectively. We can assume that $I=[a,b]$ where $a,b\in\mathbb{R}$.

First, we want to prove $\mu^*(I)\leq\ell(I)$. Let $\epsilon>0$, we have

$$I\subseteq(a,b)\cup(a-\epsilon,a+\epsilon)\cup(b-\epsilon,b+\epsilon).$$

Thus,

$$\mu^*(I)\leq\ell(a,b)+\ell(a-\epsilon,a+\epsilon)+\ell(b-\epsilon,b+\epsilon)=(b-a)+2\epsilon+2\epsilon=b-a+4\epsilon.$$

As $\epsilon>0$ is arbitrary, we conclude that $\mu^*(I)\leq b-a=\ell(I)$.

Then, we want to prove that $\mu^*(I)\geq\ell(I)$. Let $\{I_k\}$ be any sequence of open intervals that covers I. Since I is compact, by the Heine-Borel theorem, there is a finite subcollection $\{J_i:1\leq i\leq n\}$ of I_k that still covers I. By reordering and deleting if necessary, we can assume that

$$a\in J_1=(a_1,b_1),\ b_1\in J_2=(a_2,b_2),\ ...,\ b_{n-1}\in J_n=(a_n,b_n),$$

where $b_{n-1}\leq b<b_n$. We then can compute that

$$b-a<b_n-a_1=\sum_{i=2}^{n}(b_i-b_{i-1})+(b_1-a_1)<\sum_{i=1}^{n}\ell(J_i)\leq\sum_{i=1}^{\infty}\ell(I_k).$$

Therefore, $\ell(I)\leq\mu^*(I)$. We can now conclude that $\mu^*(I)=\ell(I)$. This proves the result for closed and bounded intervals.

Suppose that $I = (a, b)$ is an open and bounded interval. Then, $\mu^*(I) \leq \ell(I)$ as above and

$$b - a = \mu^*([a, b]) \leq \mu^*((a, b)) + \mu^*(a) + \mu^*(b) = \mu^*((a, b)).$$

Hence $\ell(I) \leq \mu^*(I)$. The proof for half-open intervals is similar.

Finally, suppose that I is an infinite interval and let $M > 0$. There exists a bounded interval $J \subseteq I$ such that $\mu^*(J) = \ell(J) = M$ and it follows that $\mu^*(I) \geq \mu^*(J) = M$. Since $M > 0$ was arbitrary, $\mu^* = \infty = \ell(I)$. This completes the proof.

Example 2. *What is the outer measure of the set of irrational numbers in the interval [0,1]?*

Let A be the set of irrational numbers in $[0, 1]$. Since $A \subseteq [0, 1]$, then

$$\mu^*(A) \leq 1.$$

Let Q be the set of rational numbers in $[0, 1]$. Note that $[0, 1] = A \cup Q$. By Theorem 2.1 part (e) and (f), we can conclude that

$$1 \leq \mu^*(A) + \mu^*(Q).$$

Since Q is countable, then by Theorem 2.1 part (b), $\mu^*(Q) = 0$. Therefore, $\mu^*(A) = 1$.

$\square$

2.2 Measure

Definition 2.2. A set $E \subseteq \mathbb{R}$ is Lebesgue measurable if for each set $A \subseteq \mathbb{R}$, the equality $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap \overline{E})$ is satisfied. If E is a Lebesgue measurable set, then the Lebesgue measure of E is its Lebesgue outer measure and will be written as $\mu(E)$.

Since the Lebesgue outer measure satisfies the property of subadditivity, then we always have $\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \cap \overline{E})$[1] and we only need to check the reverse inequality.

Note that there is always a set E that can divide A into two mutually exclusive sets, $A \cap E$ and $A \cap \overline{E}$. But only when $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap \overline{E})$ holds, the set E is Lebesgue measurable. The latter theorem will show some properties of measurable sets.

Theorem 2.2. *The collection of measurable sets defined on $\mathbb{R}$ has the following properties:*

a) Both $\emptyset$ and $\mathbb{R}$ are measurable.

b) If E is measurable, then $\overline{E}$ is measurable.

c) If $\mu^*(E) = 0$, then E is measurable.

d) If E_1 and E_2 are measurable, then $E_1 \cup E_2$ and $E_2 \cap E_2$ are measurable.

e) If E is measurable, then $E + x_0$ is measurable.

Proof. For part (a), let $A \subseteq \mathbb{R}$.

$$\begin{aligned} \mu^*(A \cap \emptyset) + \mu^*(A \cap \overline{\emptyset}) &= \mu^*(\emptyset) + \mu^*(A) \\ &= 0 + \mu^*(A) \\ &= \mu^*(A) \end{aligned}$$

$$\begin{aligned} \mu^*(A \cap \mathbb{R}) + \mu^*(A \cap \overline{\mathbb{R}}) &= \mu^*(A) + \mu^*(\emptyset) \\ &= \mu^*(A) + 0 \\ &= \mu^*(A) \end{aligned}$$

[1] $\overline{E} = \mathbb{R} \setminus E$

For part (b), if E is measurable, then for every set $A \subseteq \mathbb{R}$, such that $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap \overline{E})$. Then,

$$\begin{aligned}\mu^*(A \cap \overline{E}) + \mu^*(A \cap \overline{(\overline{E})}) &= \mu^*(A \cap \overline{E}) + \mu^*(A \cap E) \\ &= \mu^*(A)\end{aligned}$$

For part (c), let $A \subseteq \mathbb{R}$. Since $\mu^*(E) = 0$ and $A \cap E \subseteq E$, then $\mu^*(A \cap E) = 0$. We can obtain that

$$\begin{aligned}\mu^*(A) &\geq \mu^*(A \cap \overline{E}) \\ &= \mu^*(A \cap \overline{E}) + \mu^*(A \cap E),\end{aligned}$$

which implies that $\mu^*(A) = \mu^*(A \cap \overline{E}) + \mu^*(A \cap E)$ by Theorem 2.1 part (e).

For part (d), let $A \subseteq \mathbb{R}$. Note that

$$A \cap (E_1 \cup E_2) = (A \cap E_1) \cup (A \cap E_2) = (A \cap E_1) \cup (A \cap \overline{E_1} \cap E_2)$$

Then, by DeMorgan's Law and Theorem 2.1 part (e), we know that

$$\begin{aligned}\mu^*(A) &= \mu^*(A \cap E_1) + \mu^*(A \cap \overline{E_1}) \\ &= \mu^*(A \cap E_1) + \mu^*(A \cap \overline{E_1} \cap E_2) + \mu^*(A \cap \overline{E_1} \cap \overline{E_2}) \\ &\geq \mu^*(A \cap (E_1 \cup E_2)) + \mu^*(A \cap (\overline{E_1 \cup E_2})),\end{aligned}$$

showing that $E_1 \cup E_2$ is measurable. Since $E_1 \cap E_2 = \overline{(\overline{E_1} \cup \overline{E_2})}$, then the set $E_1 \cap E_2$ is measurable by Theorem 2.2 part (b).

For part (e), let $A \subseteq \mathbb{R}$. Then,

$$\begin{aligned}\mu^*(A) &= \mu^*(A - x_0) \\ &= \mu^*((A - x_0) \cap E) + \mu^*((A - x_0) \cap \overline{E}) \\ &= \mu^*(((A - x_0) \cap E) + x_0) + \mu^*(((A - x_0) \cap \overline{E}) + x_0) \\ &= \mu^*(A \cap (E + x_0)) + \mu^*(A \cap (\overline{E} + x_0))\end{aligned}$$

Therefore, $E + x_0$ is measurable. $\square$

Now let us have a look at the following example.

Example 3. *Suppose that E has measure zero where $E \subseteq \mathbb{R}$. Prove that the set $E^2 = \{x^2 : x \in E\}$ has measure zero.*

Proof. Let $E_n = E \cap (-n, n) \subseteq E$. Then,

$$E_n^2 = E^2 \cap (0, n^2) \subseteq E^2 \quad \text{and} \quad E^2 = \bigcup_{n=1}^{\infty} E_n^2.$$

Since E has measure zero, then E_n has measure zero. Let $\epsilon > 0$. Suppose there exists a sequence of intervals (a_k, b_k) such that

$$E_n \subseteq \bigcup_{k=1}^{\infty} (a_k, b_k) \quad \text{and} \quad \sum_{k=1}^{\infty} |b_k - a_k| < \frac{\epsilon}{2n}.$$

With the sake of simplicity, we only consider the situation that $0 < a_k < b_k$. Since

$$\mu(a_k^2, b_k^2) = |b_k^2 - a_k^2| = |b_k + a_k| \cdot |b_k - a_k| \leq 2n|b_k - a_k|,$$

then

$$\mu(E_n^2) \leq \sum_{k=1}^{\infty} \mu(a_k^2, b_k^2) \leq \sum_{k=1}^{\infty} 2n|b_k - a_k| = \epsilon.$$

It implies that the measure of E_n^2 is zero, which shows that the measure of E^2 is zero. □

Then how should we measure the collection of disjoint measurable sets?

Lemma 2.3. *Let $E_i : 1 \leq i \leq n$ be a finite collection of disjoint measurable sets. If $A \subseteq \mathbb{R}$, then*

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

Proof. We will prove this by the principle of mathematical induction. When $n = 1$, the equality holds. Suppose that the statement is valid for $n - 1$ disjoint measurable sets when $n > 1$. Then, when there are n disjoint measurable sets,

$$\begin{aligned}
\mu^*\left(A \cap (\bigcup_{i=1}^{n} E_i)\right) &= \mu^*\left(A \cap (\bigcup_{i=1}^{n} E_i) \cap E_n\right) + \mu^*\left(A \cap (\bigcup_{i=1}^{n} E_i) \cap \overline{E_n}\right) \\
&= \mu^*(A \cap E_n) + \mu^*\left(A \cap (\bigcup_{i=1}^{n-1} E_i)\right) \\
&= \mu^*(A \cap E_n) + \sum_{i=1}^{n-1} \mu^*(A \cap E_i) \\
&= \sum_{i=1}^{n} \mu^*(A \cap E_i)
\end{aligned}$$

□

Note that when $A = \mathbb{R}$,

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

Theorem 2.4. *If $\{E_i\}$ is a sequence of disjoint measurable sets, then* $\mu\left(\bigcup\limits_{i=1}^{\infty} E_i\right) = \sum\limits_{i=1}^{\infty} \mu(E_i)$.

Proof. According to Lemma 2.3, $\sum\limits_{i=1}^{n} \mu(E_i) = \mu\left(\bigcup\limits_{i=1}^{n} E_i\right) \leq \mu\left(\bigcup\limits_{i=1}^{\infty} E_i\right)$ for each positive integer n, which implies that $\sum\limits_{i=1}^{\infty} \mu(E_i) \leq \mu\left(\bigcup\limits_{i=1}^{\infty} E_i\right)$. By countably subadditive property, $\sum\limits_{i=1}^{\infty} \mu(E_i) \geq \mu\left(\bigcup\limits_{i=1}^{\infty} E_i\right)$. Therefore, $\mu\left(\bigcup\limits_{i=1}^{\infty} E_i\right) = \sum\limits_{i=1}^{\infty} \mu(E_i)$. □

The previous theorem shows that if A and B are disjoint measurable sets, then $\mu(A \cup B) = \mu(A) + \mu(B)$. If $\{A_i\}$ is a sequence of disjoint measurable sets, then $\mu\left(\bigcup\limits_{i=1}^{\infty} A_i\right) = \sum\limits_{i=1}^{\infty} \mu(A_i)$. As so far, we have already seen that when the sets are measurable, Lebesgue measure satisfies property (1),(2),(3) and (4). But what kinds of sets are measurable? Certainly every interval is measurable.

Theorem 2.5. *Every interval is measurable.*

Proof. Suppose that $A \subseteq \mathbb{R}$. Let $a \in \mathbb{R}$, $A_1 = A \cap (-\infty, a]$ and $A_2 = A \cap (a, \infty)$. Then A_1 and A_2 are complements of each other on $\mathbb{R}$. By Theorem 2.1 part (e), we know that $\mu^*(A_1) + \mu^*(A_2) \geq \mu^*(A)$. In order to prove A_1 is a measurable set, we still need to show that $\mu^*(A_1) + \mu^*(A_2) \leq \mu^*(A)$. If $\mu^*(A) = \infty$, the previous statement is trivial. If $\mu^*(A) < \infty$, let $\epsilon > 0$. By definition, there exists a sequence of open intervals $\{I_k\}$ such that $A \subseteq \bigcup_{k=1}^{\infty} I_k$ and $\sum_{k=1}^{\infty} \ell(I_k) < \mu^*(A) + \epsilon$. Let $I_k^1 = I_k \cap (-\infty, a]$ and $I_k^2 = I_k \cap (a, \infty)$. Then $\{I_k^1\}$ and $\{I_k^2\}$ are two sequences of intervals covering set A_1 and A_2 respectively. Since

$$\mu^*(I_k^1) + \mu^*(I_k^2) = \ell(I_k^1) + \ell(I_k^2) = \ell(I),$$

then

$$\mu^*(A_1) + \mu^*(A_2) \leq \sum_{k=1}^{\infty} \ell(I_k^1) + \sum_{k=1}^{\infty} \ell(I_k^2) = \sum_{k=1}^{\infty} \ell(I_k) < \mu^*(A) + \epsilon,$$

which implies that $\mu^*(A_1) + \mu^*(A_2) \leq \mu^*(A)$. Therefore, every interval is measurable.

□

Theorem 2.6. *If E_i is a sequence of measurable sets, then $\bigcup_{i=1}^{\infty} E_i$ and $\bigcap_{i=1}^{\infty} E_i$ are measurable sets.*

Proof. Let $E = \bigcup_{i=1}^{\infty} E_i$, let $H_1 = E_1$, and let $H_n = E_n - \bigcup_{i=1}^{n-1} E_i$ for each $n \geq 2$. Then $\{H_n\}$ is a sequence of disjoint measurable sets and $E = \bigcup_{n=1}^{\infty} H_n$. Note that $\overline{E} \subseteq \overline{\bigcup_{i=1}^{n} H_i}$ for each n. Let $A \subseteq \mathbb{R}$, then using the previous lemma,

$$\mu^*(A) = \mu^*\left(A \cap (\bigcup_{i=1}^{n} H_i)\right) + \mu^*\left(A \cap \overline{(\bigcup_{i=1}^{n} H_i)}\right) \geq \sum_{i=1}^{n} \mu^*(A \cap H_i) + \mu^*(A \cap \overline{E})$$

for each n.

Since $A \cap E = \bigcup_{i=1}^{\infty} (A \cap H_i)$, then by the countable subaddition property,

$$\begin{aligned} \mu^*(A \cap E) + \mu^*(A \cap \overline{E}) &= \mu^* \left(\bigcup_{i=1}^{\infty} (A \cap H_i) \right) + \mu^*(A \cap \overline{E}) \\ &\leq \sum_{i=1}^{\infty} \mu^*(A \cap H_i) + \mu^*(A \cap \overline{E}) \\ &\leq \mu^*(A) \end{aligned}$$

Since $\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \cap \overline{E})$ according to countable subaddition property again, then we can conclude that E is a measurable set. Since

$$\bigcap_{i=1}^{\infty} E_i = \overline{(\bigcup_{i=1}^{\infty} \overline{E_i})},$$

the set $\bigcap_{i=1}^{\infty} E_i$ is measurable by the first part of the proof. □

Theorem 2.7. *Every open set and every closed set is measurable.*

Every interval is measurable and their intersection and union are also measurable by Theorem 2.6, so we can see that every open set and every closed set is measurable.

Since all the open and closed sets are measurable, and their finite collection is closed under countable intersections and unions, then it is really hard to imagine that there exists a set that is not measurable. The following example will show that there are non-measurable sets.

Theorem 2.8. *There exist sets that are not measurable.*

Proof. Define an equivalence relation $\sim$ on $\mathbb{R}$ by $x \sim y$ if $x - y$ is rational. This relation establishes a collection of equivalence classes of the form $\{x+r : r \in \mathbb{Q}\}$. Each equivalence class contains a point in $[0,1]$. Let $E \subseteq [0,1]$ be a set that consists of one point from each equivalence class. Let $\{r_i\} = \mathbb{Q} \cap [-1,1]$ and let $E_i = E + r_i$ for each i.

Next, we want to show that

$$[0,1] \subseteq E_i \subseteq [-1,2].$$

The second part is straightforward. To prove the first, let $x \in [0,1]$. There exists $y \in E$ such that $x - y$ is rational. Since $-1 \leq x - y \leq 1$, there exists an index j such that $x - y = r_j$. Therefore, $x = y + r_j \in E_j$. Furthermore, the sets E_i are disjoint. Otherwise, there exists $y, z \in E$ such that $y + r_i = z + r_j$, which implies that $y \sim z$, a contradiction.

Now suppose that E is a measurable set. Then each E_i is measurable and $\mu(E_i) = \mu(E)$, since $E_i = E + r_i$. Then, we have the following statement

$$1 \leq \mu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E) \leq 3,$$

a contradiction.

Therefore, E is not measurable. $\square$

The above proof only uses the property (1),(2),(3) and (4), which implies that there is no measure satisfies all of the 4 properties on all subsets of $\mathbb{R}$. Consider the following statement $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$ for all disjoint sets A and B. Using the statement from the previous proof, we have

$$n\mu^*(E) = \sum_{i=1}^{n} \mu^*(E) = \mu^*(\bigcup_{i=1}^{n} E_i) \leq \ell([-1,2]) = 3, \quad \text{for all } n.$$

The only possible value for n is $n = 0$, i.e $\mu^*(E) = 0$, but then E would be a measurable set, a contradiction. Therefore, the statement $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$ only holds for disjoint measurable sets. And that is the reason why we only consider measurable sets, which brings us to the next two definitions.

Definition 2.3. A collection $\mathcal{A}$ of sets is an **algebra** if $\emptyset \in \mathcal{A}$, $\overline{E} \in \mathcal{A}$ whenever $E \in \mathcal{A}$, and $\mathcal{A}$ is closed under finite unions (and hence finite intersections).

Definition 2.4. An algebra $\mathcal{A}$ is a **σ-algebra** if $\mathcal{A}$ is closed under countable unions (and hence countable intersections).

The collection of all subsets of $\mathbb{R}$ is a σ-algebra and the collection of Lebesgue measurable sets is a σ-algebra. Also, the intersection of σ-algebras is a σ-algebra.

Let $\mathcal{B}$ be the intersection of all σ-algebras that contain the open sets. Then $\mathcal{B}$ is the smallest σ-algebra that contains all the open sets. Any set in $\mathcal{B}$ is called a Borel set. All countable sets, intervals, open sets and closed sets are Borel sets.

The two most special and frequently used classes of Borel set are G_δ and F_σ. A G_δ set is any set that can be expressed as a countable intersection of open sets. An F_σ set is any set that can be expressed as a countable union of closed sets.

Here are some properties about these two sets.

(1) The complement of an F_σ set is a G_δ set.

(2) The complement of a G_δ set is an F_σ set.

(3) The union of countably many F_σ sets is an F_σ set.

(4) The intersection of countably many G_δ sets is a G_δ set.

(5) The intersection of two F_σ sets is an F_σ set.

(6) The union of two G_δ sets is a G_δ set.

(7) Every open set is an F_σ set.

(8) Every closed set is a G_δ set.

(9) The set difference of two closed sets is an F_σ set.

Note that each even-numbered property can be proved by taking complement of previous odd-numbered property. So we will focus on the odd-numbered property.

To prove (1), let F be an F_σ set. Then, by definition, there exists a set of countable closed sets C_i such that $F = \bigcup\limits_{i=1}^{\infty} C_i$. Let $G = \mathbb{R} \setminus F$. Then, by De Morgan's Laws,

$$G = \mathbb{R} \setminus F = \mathbb{R} \setminus \bigcup_{i=1}^{\infty} C_i = \bigcap_{i=1}^{\infty} (\mathbb{R} \setminus C_i).$$

Since C_i are closed sets, then $(\mathbb{R} \setminus C_i)$ are open sets, which implies that G is a G_δ set.

Property (3) can be shown directly by the definition of union. The union of countably many F_σ sets is still a countable union of closed sets, which is an F_σ set.

For property (5), the intersection of two F_σ sets is an intersection of two countable unions of closed sets, which will generates a countable union of closed sets and it is an F_σ. Suppose that F_i and F_j are two F_σ sets and $F_i \cap F_j = F$. By definition, there exists two sets of countable closed sets C_i and C_j such that

$$F_i = \bigcup_{i=1}^{\infty} C_i \quad \text{and} \quad F_j = \bigcup_{j=1}^{\infty} C_j.$$

Then,

$$\begin{aligned} F = F_{\sigma_1} \cap F_{\sigma_2} &= \left(\bigcup_{i=1}^{\infty} C_i\right) \cap \left(\bigcup_{j=1}^{\infty} C_j\right) \\ &= \bigcup_{i=1}^{\infty} \left(C_i \cap \bigcup_{j=1}^{\infty} C_{ij}\right) = \bigcup_{i=1}^{\infty} \left(C_i \cap \bigcup_{j=1}^{\infty} C_j\right) \\ &= \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} \left(C_i \cap C_j\right), \end{aligned}$$

which implies that F is an F_σ set.

Property (7) can be shown by considering every open set can be covered by countable many closed intervals. Therefore, every open set is an F_σ set. Let F be a open set. Then there exists a sequence of open intervals (a_i, b_i) for $i = 1, 2, \ldots,$ such that

$$F = \bigcup_{i=1}^{\infty} (a_i, b_i) = \bigcup_{i=1}^{\infty} [a_i + \epsilon, b_i - \epsilon].$$

Since ϵ is arbitrary, then F is an F_σ set.

For property (9), we can prove it by definition. Suppose A and B are two closed sets. Then

$$A \setminus B = A \cap \overline{B}.$$

The set A can be expressed as a countable union of closed set, thus it is a F_σ set. B is a closed set, then $\overline{B}$ is an open set, which is also an F_σ. Therefore, the intersection of two F_σ sets is an F_σ set.

The next theorem will show us how these two special Borel sets coordinate with measurable sets. It also provides us a different perspective to define a measurable set.

Theorem 2.9. *For any set $E \subseteq R$, the following statements are equivalent:*

(1) The set E is measurable.

(2) For each $\epsilon > 0$, there is an open set O such that $E \subseteq O$ and $\mu^*(O - E) < \epsilon$.

(3) For each $\epsilon > 0$, there is a closed set K such that $K \subseteq E$ and $\mu^*(E - K) < \epsilon$.

(4) There exists a G_δ set G such that $E \subseteq G$ and $\mu^*(G - E) = 0$.

(5) There exists an F_σ set F such that $F \subseteq E$ and $\mu^*(E - F) = 0$.

Proof. The best approach is to prove (1) $\to$ (2) $\to$ (4) $\to$ (1) $\to$ (3) $\to$ (5) $\to$ (1)

Let E be a measurable set. We will consider the following two situations respectively, $\mu(E) < \infty$ and $\mu(E) = \infty$. Suppose that $\mu(E) < \infty$ and let $\epsilon > 0$. There is a sequence $\{I_k\}$ of open intervals such that $E \subseteq \bigcup\limits_{k=1}^{\infty} I_k$ and $\sum\limits_{k=1}^{\infty} \ell(I_k) < \mu(E) + \epsilon$. Let $O = \bigcup\limits_{k=1}^{\infty} I_k$. Since $O = E \cup (O - E)$, which is a union of disjoint measurable sets, then we have

$$\mu(O - E) = \mu(O) - \mu(E) \leq \sum_{k=1}^{\infty} \ell(I_k) - \mu(E) < \epsilon.$$

Suppose that $\mu(E) = \infty$ and let $E_n = E \cap \{x : n-1 \leq |x| < n\}$ for each positive integer n. Let $\epsilon > 0$. As above, for each positive integer n, there exists an open set O_n such that $E_n \subseteq O_n$ and $\mu(O_n - E_n) < \frac{\epsilon}{2^n}$. Note that the set $O = \bigcup\limits_{n=1}^{\infty} O_n$

is open and $E \subseteq O$. Since $O - E \subseteq \bigcup_{n=1}^{\infty}(O_n - E_n)$, then

$$\mu(O - E) \leq \sum_{n=1}^{\infty} \mu(O_n - E_n) < \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon.$$

Therefore, we proved (1) $\rightarrow$ (2).

Suppose that (2) is true. For each positive integer n, there exists an open set O_n such that $E \subseteq O_n$ and $\mu^*(O_n - E) < \frac{1}{n}$. Let $G = \bigcap_{n=1}^{\infty} O_n$. Then G is a G_δ set, $E \subseteq G$, and

$$\mu^*(G - E) \leq \mu^*(O_n - E) < \frac{1}{n},$$

for all n. Therefore, $\mu^*(G - E) = 0$, showing that (2) $\rightarrow$ (4)

Suppose that (4) is true. Since $\mu^*(G - E) = 0$, the set $G - E$ is measurable. It follows that $E = G \cap \overline{(G - E)}$ is the intersection of two measurable sets, which implies that E is measurable. Therefore, we proved (4) $\rightarrow$ (1).

Note that the proof of (1) $\rightarrow$ (3) follows the proof of (1) $\rightarrow$ (2) by taking complements. And the rest of the proof are similar to the previous proof, so will be omitted. □

Next, we will introduce some new ideas about sets.

Definition 2.5. Let E be a subset of $\mathbb{R}$, then

(1) The set E is **perfect** if E is closed and each point of E is a limit point[1] of E.

(2) The set E is **nowhere dense** if $\overline{E}$ (the closure of E) contains no open intervals.

[1] x_0 is a limit point of E if every deleted neighborhood of x_0 contains a point of E. A deleted neighborhood of a point x_0 is a set that contains every point of some neighborhood of x_0 except for x_0 itself. For example, $E = \{x : 0 < |x - x_0| < \epsilon\}$ is a deleted neighborhood of x_0.

(3) The set E is **dense** if every open interval (a, b) contains a member of E. Or a set E is called **dense** in $\mathbb{R}$ if $\overline{E} = \mathbb{R}$

Now let us see some examples to understand the first two ideas better.

Example 4. *Let both set A and set B be subsets of $\mathbb{R}$, and we define $A = [0, 1]$ and $B = \mathbb{Z}$. Set A is a perfect set but is not nowhere dense. Set B is nowhere dense in $\mathbb{R}$, but is not a perfect set, because it has no limit points.*

In 1883, German mathematician Georg Cantor introduced an idea that combined perfect and nowhere dense[4]. It is called the Cantor set.

Definition 2.6. The **Cantor Set** is a closed subset of $[0, 1]$. Let

$$\begin{aligned}
G_1 &= \left(\frac{1}{3}, \frac{2}{3}\right), \\
G_2 &= \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right), \\
G_3 &= \left(\frac{1}{27}, \frac{2}{27}\right) \cup \cdots \cup \left(\frac{25}{27}, \frac{26}{27}\right), \text{ and etc.}
\end{aligned}$$

Let $G = \bigcup_{k=1}^{\infty} G_k$, so G is an open subset of $(0, 1)$. Then, we define the Cantor set

$$C = [0, 1] \setminus G.$$

Note that $\mu(G_1) = \frac{1}{3}$, $\mu(G_2) = \frac{2}{9}$, $\mu(G_3) = \frac{4}{27}$ and etc. Thus,

$$\begin{aligned}
\mu(G) &= \frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \dots \\
&= \frac{1}{3}\left(1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \dots\right) \\
&= \frac{1}{3} \cdot \frac{1}{1 - \frac{2}{3}} \\
&= 1
\end{aligned}$$

Therefore, $\mu(C) = 0$.

Theorem 2.10. *The Cantor set C is a nonempty, perfect, nowhere dense set of measure zero.*

Proof. Since $\mu(C) = 0$, it does not contain any intervals, which implies that the set C is nowhere dense. In order to prove that set C is perfect, we need to show that every point in C is a limit point of C. Let $x \in C$ and $\delta > 0$. Choose an integer n such that $3^{-n} < \delta$. Since $x \in K_n$, then there exists a closed interval I of length 3^{-n} such that $x \in I \subseteq K_n$. Let a be an end point of I that is distinct from x and note that $a \in C$ and $0 < |x - a| < \delta$. Therefore, x is a limit point of set C, showing that C is perfect. □

One of the functions that derives from the Cantor set is called the Cantor function. It is a counterexample for a number of conjectures in the theory of Lebesgue integration [5]. We will construct the Cantor function in the following steps.

First, we define $\{K_n\}$ as a sequence of closed sets as follows:

$$\begin{aligned}
K_1 &= \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right], \\
K_2 &= \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{3}{9}\right] \cup \left[\frac{6}{9}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right], \\
K_3 &= \left[0, \frac{1}{27}\right] \cup \left[\frac{2}{27}, \frac{3}{27}\right] \cup \left[\frac{6}{27}, \frac{7}{27}\right] \cup \left[\frac{8}{27}, \frac{9}{27}\right] \cup \left[\frac{18}{27}, \frac{19}{27}\right] \cup \left[\frac{20}{27}, \frac{21}{27}\right] \cup \\
&\quad \left[\frac{24}{27}, \frac{25}{27}\right] \cup \left[\frac{26}{27}, 1\right].
\end{aligned}$$

The set K_n is the union of 2^n disjoint closed intervals and each interval has length is 3^{-n}. Then for each n, let E_n be the closure of the set $([0, 1] - K_n) \cup \{0, 1\}$. Then we have followings:

$$\begin{aligned}
E_1 &= \{0\} \cup [\frac{1}{3}, \frac{2}{3}] \cup \{1\}, \\
E_2 &= \{0\} \cup [\frac{1}{9}, \frac{2}{9}] \cup [\frac{1}{3}, \frac{2}{3}] \cup [\frac{7}{9}, \frac{8}{9}] \cup \{1\}, \text{and etc.}
\end{aligned}$$

For each n, let f_n be the nondecreasing, continuous function defined on $[0, 1]$ as follows:

(1) $f_n(0) = 0$ and $f_n(1) = 1$;

(2) f_n is constant on each subinterval of E_n and takes on the values $1/2^n$, $2/2^n$,...,$(2^n - 1)/2^n$ in increasing order;

(3) f_n is linear on the intervals forming the complement of E_n.

Note that $f_m = f_n$ on E_n for all $m > n$ and that f_n is constant on the intervals comprising E_n. The limit f of this sequence of functions is known as the Cantor function.

Theorem 2.11. *The Cantor function f is continuous and nondecreasing on $[0,1]$. Furthermore, the function f is constant on each interval of $[0,1]-C$ and $f(C) = [0,1]$.*

Proof. It is not difficult to show that $|f_{n+1}(x) - f_n(x)| \leq 2^{-n}$ for all n and for all $x \in [0,1]$. Furthermore, the function f is constant on each interval of $[0,1]-C$ and $f(C) = [0,1]$. □

Here are some properties of the Cantor function, named C. Then,

- C is defined everywhere in the interval $[0,1]$.
- C is continuous and increasing but not absolutely continuous.
- C is differentiable almost everywhere in the interval $[0,1]$
- C is constant on each interval, but not constant overall.
- C maps the Cantor set C onto $[0,1]$.

We will stop here for this section.

2.3 Inner Measure

In previous sections, we have already discussed Lebesgue outer measure and Lebesgue measure. Intuitively, there should be another measure named Lebesgue inner measure. Let's define the inner measure and see some basic properties.

Definition 2.7. Let E be a subset of $\mathbb{R}$. The inner measure of E is defined by

$$\mu_*(E) = \sup\{\mu(K) : K \subseteq E \text{ and } K \text{ is closed}\}.$$

Recall that Lebesgue outer measure of a set E uses an infimum of the union of a sequence open sets that cover the set E, while Lebesgue inner measure of a set E uses a supremum of a set inside the set E. Then, it is obvious that $\mu_*(E) \le \mu^*(E)$ for any set E. Also, for $A \subseteq B$, $\mu_*(A) \le \mu^*(B)$.

Theorem 2.12. *Let A and E be subsets of $\mathbb{R}$.*

(1) Suppose that $\mu^*(E) < \infty$. Then E is measurable if and only if $\mu_*(E) = \mu^*(E)$.

(2) If E is measurable and $A \subseteq E$, then $\mu(E) = \mu_*(A) + \mu^*(E - A)$.

Proof. For part (1), suppose that E is a measurable set and let $\epsilon > 0$. According to Theorem 2.9, there exists a closed set K such that $K \subseteq E$ and $\mu(E - K) < \epsilon$. Thus,

$$\mu^*(E) \ge \mu_*(E) \ge \mu(K) > \mu(E) - \epsilon = \mu^*(E) - \epsilon,$$

which implies that the inner measure and outer measure of E are equal. Now let's prove the reverse direction. Suppose that $\mu_*(E) = \mu^*(E)$. Let $\epsilon > 0$. Then there exists a closed set K and an open set G such that $K \subseteq E \subseteq G$ and

$$\mu(K) > \mu_*(E) - \frac{\epsilon}{2} \quad \text{and} \quad \mu(G) < \mu^*(E) + \frac{\epsilon}{2}.$$

Then we find that

$$\mu^*(G - E) \le \mu^*(G - K) = \mu(G - K) = \mu(G) - \mu(K) < \epsilon.$$

According to Theorem 2.9, the set E is measurable.

For part (2), let $\epsilon > 0$. There exists a closed set $K \subseteq A$ such that $\mu(K) > \mu_*(A) - \epsilon$. Then,

$$\mu(E) = \mu(K) + \mu(E - K) > \mu_*(A) - \epsilon + \mu^*(E - A)$$

and it follows that $\mu(E) \geq \mu_*(A) + \mu^*(E - A)$. According to Theorem 2.9, there exists a measurable set B such that $E - A \subseteq B \subseteq E$ and $\mu(B) = \mu^*(E - A)$. Since $E - B \subseteq A$, it follows that $\mu_*(E - B) \geq \mu_*(A)$. Thus,

$$\mu(E) = \mu(B) + \mu(E - B) = \mu^*(E - A) + \mu(E - B) \leq \mu^*(E - A) + \mu_*(A)$$

By combining the two inequalities, we can obtain $\mu(E) = \mu_*(A) + \mu^*(E - A)$. $\square$

We have presented three basic measures: Lebesgue outer measure, Lebesgue measure, and Lebesgue inner measure. These three measures are different from each other but also coordinate with each other. For any generic sets, we can have its Lebesgue outer(inner) measure but we cannot always do the Lebesgue measure on that sets. We need to identify whether such a set is measurable or not.

One important application of measure theory is probability theory. In everyday conversation, the probability is a measure of one's belief in the occurrence of a future event (see [6]). We can regard an "event" in probability theory as a "set" in measure theory. Then the probability of such a event will be the measure of a corresponding set.

In next chapter, we will define the Lebesgue Integral and compare it the Riemann Integral.

3 Lebesgue Integral

3.1 Lebesgue Integral

According to "Interactive Real Analysis" organized by Bert G. Wachsmuth and Paul Golba (see [7]), we construct the Riemann integral in the following steps:

- subdivide the domain of the function (usually a closed, bounded interval) into finitely many subintervals (the partition)
- construct a step function that has a constant value on each of the subintervals of the partition (the Upper and Lower sums)
- take the limit of these step functions as you add more and more points to the partition.

The "opposite" approach, named the Lebesgue integral, will be operated in the following steps:

- subdivide the range of the function into finitely many intervals
- construct a simple function by taking a function whose values are those finitely many numbers
- take the limit of these simple functions as you add more and more points in the range of the original function

First, in particular, we need a function that can help us distinguish whether a given value x is in the measurable set A_i. We call this function the characteristic function. The following statement is the formal definition of characteristic function and introduces the simple function.

Definition 3.1. For any set A, the function

$$\mathcal{X}_A(x) = \begin{cases} 1, & x \in A, \\ 0, & \text{otherwise.} \end{cases}$$

is called the characteristic function of set A. A linear combination of characteristic functions,

$$\varphi(x) = \sum_{i=1}^{n} a_i \mathcal{X}_{A_i}(x),$$

is called a simple function if the sets A_i are measurable.

For a function f defined on a measurable set A that takes no more than finitely many distinct values ($i.e.\, a_1, ..., a_n$), the function f can always be written as a simple function

$$f(x) = \sum_{i=1}^{n} a_i \mathcal{X}_{A_i}(x),$$

where $A_i = \{x \in A : f(x) = a_i\}$. Therefore, simple functions can be thought of as dividing the range of f, where resulting sets A_i may or may not be intervals[7].

Let us pause for a second. We want to ask ourselves: is the simple function $\varphi(x)$ unique? The answer is no. Because we might define different disjoint sets that have a same function value. The simplest expression is

$$\varphi(x) = \sum_{i=1}^{n} a_i \mathcal{X}_{A_i}(x),$$

where $A_i = \{x \in A : \varphi(x) = a_i\}$. At this time, the constants a_i are distinct, the sets A_i are disjoint and we call that representation the canonical representation of ϕ.

Then, for simple functions, we define the Lebesgue integral as follows:

Definition 3.2. If $\varphi(x) = \sum\limits_{i=1}^{n} a_i \mathcal{X}_{A_i}(x)$ is a simple function and $\mu(A_i)$ is finite for all i, then the Lebesgue integral of $\varphi(x)$ is defined as

$$\int_E \varphi(x)\, dx = \sum_{i=1}^{n} a_i \mu(A_i).$$

Definition 3.3. Suppose f is a bounded function defined on a measurable set E with finite measure. We define the upper and lower Lebesgue integrals, respectively, as

$$I^*(f)_L = \int_E \inf\{\varphi(x)\, dx : \varphi \text{ is simple and } \varphi \geq f\},$$

$$I_*(f)_L = \int_E \sup\{\varphi(x)\, dx : \varphi \text{ is simple and } \varphi \leq f\}.$$

If $I^*(f)_L = I_*(f)_L$, then the function f is called Lebesgue integrable over set E and the Lebesgue integral of f over set E is denoted by

$$\int_E f(x)\, dx.$$

3.2 Riemann vs. Lebesgue

The previous two sections recalled the Riemann integral and introduced the Lebesgue integral. We realize that both of them can help us to integrate functions. The difference is that the Riemann integral subdivides the domain of a function, while the Lebesgue integral subdivides the range of that function. The step function for Riemann integral has a constant value in each of the subintervals of the partition, while the simple function for Lebesgue integral provides finitely many measurable sets corresponding to each value of that function.

The improvement from the Riemann integral to the Lebesgue integral is that the Lebesgue integral provides more generality than the Riemann integral does. From the reverse perspective, the Riemann integral can imply the Lebesgue integral.

Theorem 3.1. *If f is a bounded function defined on $[a,b]$ such that f is Riemann integrable, then f is Lebesgue integrable and*

$$(R)\int\limits_a^b f(x)\,dx = \int\limits_{[a,b]} f(x)\,dx$$

Proof. We mentioned in Section 3.1 that for a given function f, we defined

$$(R)\overline{\int_a^b} f(x)\,dx = \inf S \quad \text{and} \quad (R)\underline{\int_a^b} f(x)\,dx = \sup s.$$

Since every step function is a simple function, then every upper sum is greater than f and every lower sum is less than f. Therefore, we can have the following inequalities

$$(R)\underline{\int_a^b} f(x)\,dx \leq \sup_{\varphi \leq f} \int\limits_a^b \varphi(x)\,dx \leq \inf_{\psi \geq f} \int\limits_a^b \psi(x)\,dx \leq (R)\overline{\int_a^b} f(x)\,dx.$$

Since f is Riemann integrable, the first and last quantities are equal to each other, then by Definition 3.3, f is Lebesgue integrable on $[a,b]$ $\square$

Example 5. *Recall the function:*

$$f(x) = \begin{cases} 0, & x \text{ is rational.} \\ 1, & x \text{ is irrational.} \end{cases}$$

We define this function $f(x)$ on $[0,1]$. We will find the Lebesgue integral of this function.

Its Lebesgue integral is given by

$$\int_{[0,1]} f = 0 \cdot |A| + 1 \cdot |B|$$

where $A = [0,1] \cap \mathbb{Q}$ is the set of rational numbers in $[0,1]$ and $B = [0,1] \setminus \mathbb{Q}$ is the set of irrational numbers, and μ denotes the Lebesgue measure of a set. Recall that the Lebesgue measure of a set is a generalization of the length of an interval which applies to more general sets. It turns out that $\mu(A) = 0$ and $\mu(B) = 1$. Thus, the Lebesgue integral of this function is 1.

3.3 The Measurable Function

A thoughtful reader must have already noticed that we restricted our Lebesgue integrable function to bounded functions. Can we extend teh Lebesgue integral to unbounded functions? To answer this question, we need to define a measurable function.

Definition 3.4. Let f be a function from $E \subset R$ into $R \cup (-\infty, \infty)$. The function f is (Lebesgue) measurable if

- the domain E of the function is a measurable
- for every real number a the set $f^{-1}(-\infty, a)$ is a measurable set.

The following theorem follows from our previous assumption.

Theorem 3.2. *If f is a bounded function defined on a measurable set E with finite measure. Then f is measurable if and only if f is Lebesgue integrable.*

Without the restriction of bounded function, this theorem does not hold. A function can be measurable but not Lebesgue integrable. Consider the following function defined on $(0, 1]$

$$f(x) = \frac{1}{x}.$$

Theorem 3.3. *Every continuous function is measurable.*

Proof. Let f^{-1} is defined on a set O. Since f is continuous, then $f^{-1}(O)$ is open if O is open. Since open sets are measurable, then the function is measurable followed by the definition. $\square$

References

[1] William F. Trench. *Introduction to real analysis.* s.n., 2009.

[2] (Riemann) Integration Sucks!!! https://math.berkeley.edu/ peyam/Math1AFa10/ Integration%20Sucks!!!.pdf. Accessed: 2017-05-10.

[3] Charles Chapman. Pugh. *Real mathematical analysis.* Undergraduate Texts in Mathematics. Springer, 2004.

[4] Georg Cantor. *ber unendliche, lineare Punktmannigfaltigkeiten V" [On infinite, linear point-manifolds (sets)]*, volume 21. Mathematische Annalen, 1883.

[5] Russell A. Gordon. *The integrals of Lebesgue, Denjoy, Perron, and Henstock*, volume 4. Ameri. Math. Soc., 1994.

[6] Dennis D. Wackerly, William Mendenhall, and Richard L. Scheaffer. *Mathematical statistics with applications.* Brooks-Cole, 2008.

[7] Interactive real analysis. http://www.mathcs.org/analysis/reals/. Accessed: 2017-05-10.