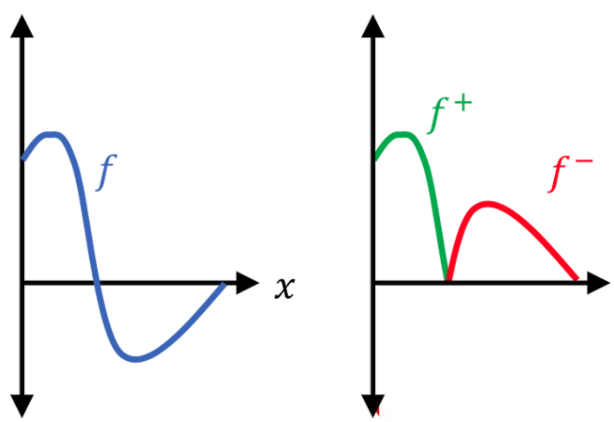
An Introduction to Measure Theory and Real Analysis

Fall 2024, Math 454 Course Notes

McGill University



Jiajun Zhang

January 17, 2025

Acknowledgments

I would like to extend my deepest thanks and appreciation to the following people, without whose support this note would not have been possible:

[Linan Chen], *Associate Professor, McGill University*
I would like to express my deepest gratitude to Professor Chen, the instructor for this course, whose guidance and expertise were invaluable throughout the development of my notes.

[Tianyi Xu], *Undergraduate, Honors Computer Science and Mathematics, McGill University*
I would like to express my deepest gratitude to my friend Tianyi Xu, who examined this notes thoughtfully and gave an enormous amount of useful suggestions.

Despite all efforts, there may still be some typos, unclear explanations, etc. If you find potential mistakes, or any suggestions regarding concepts or formats, etc., feel free to reach out to the author at *zhangjohnson729@gmail.com*.

Contents

Chapter 1

Measure Theory

σ-Algebra

Let's recap the definition of a Riemann Integral, suppose a function $f(x)$ is continuous on the interval $[a,b]$, then $\int_a^b f(x)dx$ represents the area of the region under the graph of $f(x)$ on the interval $[a,b]$. To compute this, consider a following partition:

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

Then, the Upper Riemann Integral is given by

$$\overline{\int_a^b} f(x)dx = \inf\left\{\sum_{i=1}^{n}\left(\sup_{[x_{i-1},x_i]} f(x)\right)(x_i - x_{i-1})\right\}, \text{where} \max|x_i - x_{i-1}| \to 0$$

Similarly, the Lower Riemann Integral is given by

$$\underline{\int_a^b} f(x)dx = \sup\left\{\sum_{i=1}^{n}\left(\inf_{[x_{i-1},x_i]} f(x)\right)(x_i - x_{i-1})\right\}, \text{where} \max|x_i - x_{i-1}| \to 0$$

By definition, we know that

$$\underline{\int_a^b} f(x)dx \le \int_a^b f(x)dx \le \overline{\int_a^b} f(x)dx$$

Theorem

Theorem 1. *A function $f(x)$ is said to be Riemann-Integrable over $[a,b]$ if*

$$\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx$$

This definition has some drawbacks. To illustrate this, let's take a look at one example:

Suppose $f:[a,b]\mapsto\{0,1\}$, such that $f(x)=\begin{cases}1, \text{if } \ x\in\mathbb{Q}\cap[a,b]\\ \\ 0, \text{if } \ x\in\mathbb{Q}^C\cap[a,b]\end{cases}$, we will notice that $f(x)$ is not Riemann-Integrable. Since $\mathbb{Q}$ is dense in $[a,b]$, it means that for each possible partitions, we have $\overline{\int_a^b} f(x)dx = 1\cdot(b-a)$; Similarly we can also conclude $\underline{\int_a^b} f(x)dx = 0$, thus $f(x)$ is not Riemann-Integrable.

Now, we would like to extend the notation of "integral" to include such function. Let's consider "slicing" the region horizontally, say $y_1, y_2, \cdots, y_n \in f(x)$.

Denote

$$A_i = \{x\in[a,b]; y_{i-1}\le f(x) < y_i\}$$

Another drawback arises. See that the length for each interval A_i is not fixed, and it varies according to the behavior of $f(x)$. That's why we would like to introduce the concept of a measure space.

Definition

Definition 1. *Let X be a space (i.e a non-empty set) and $\mathscr{F}$ be a collection of subsets of X (here $\mathscr{F}$ is the collection of subsets of X which we are going to measure). $\mathscr{F}$ is called a σ-algebra of subsets of X if:*

① $X\in\mathscr{F}$

② *If $A\in\mathscr{F}$, then $A^C := X\backslash A\in\mathscr{F}$. (Closed Under Taking Complement)*

③ *If a series of subsets $\{A_n, n\ge 1\}\in\mathscr{F}$, then $\bigcup_{n=1}^{\infty} A_n\in\mathscr{F}$. (Closed Under Countable Union)*

Based on the definition, the following propositions also hold.

Proposition

Proposition 1. *The definition of σ-algebra leads to:*

① $\varnothing \in \mathscr{F}$

② $X \in \mathscr{F}$

③ *If a series of subsets* $\{A_n, n \geq 1\} \in \mathscr{F}$, *then* $\bigcap_{n=1}^{\infty} A_n \in \mathscr{F}$. *(Closed Under Countable Intersection)*

④ *If* $A_1, A_2, \cdots, A_N \in \mathscr{F}$, *then* $\bigcap_{n=1}^{N} A_n \in \mathscr{F}$ *and* $\bigcup_{n=1}^{N} A_n \in \mathscr{F}$. *(Closed Under Taking Finite Union or Intersection)*

⑤ *If* $A, B \in \mathscr{F}$, *then* $A \backslash B, B \backslash A \in \mathscr{F}$

We say $\mathscr{F}_1$ is bigger than $\mathscr{F}_2$, if $\mathscr{F}_1 \subseteq \mathscr{F}_2$. In this case, for any space X, we may conclude that $\{\varnothing, X\}$ is the smallest σ-algebra and 2^X is the biggest σ-algebra. A σ-algebra can also be generated by a collection of sets.

Definition

Definition 2. *Let X be a space and $\mathscr{C}$ be a collection of subsets of X, then the σ-algebra generated by $\mathscr{C}$, denoted by $\sigma(\mathscr{C})$, is such that*

① $\sigma(\mathscr{C})$ *is itself a σ-algebra with* $\mathscr{C} \subseteq \sigma(\mathscr{C})$

② *If $\mathscr{F}'$ is a σ-algebra with $\mathscr{C} \subseteq \mathscr{F}'$, then $\sigma(\mathscr{C}) \subseteq \mathscr{F}'$. i.e $\sigma(\mathscr{C})$ is the smallest σ-algebra that is a super set of $\mathscr{C}$.*

Again, we have the following propositions:

Proposition

Proposition 2. ① $\sigma(\mathscr{C}) = \bigcap \{\mathscr{F} : \mathscr{C} \subseteq \mathscr{F}\}$ *for all σ-algebra $\mathscr{F}$*

② *If $\mathscr{C}$ itself is a σ-algebra, then* $\sigma(\mathscr{C}) = \mathscr{C}$

③ *If $\mathscr{C}_1, \mathscr{C}_2$ are 2 collections of subsets of X with $\mathscr{C}_1 \subseteq \mathscr{C}_2$, then* $\sigma(\mathscr{C}_1) \subseteq \sigma(\mathscr{C}_2)$

Borel σ-Algebra

An important example of σ-algebra on $\mathbb{R}$ (of subsets of $\mathbb{R}$) is called a *Borel σ-Algebra*.

Definition

Definition 3. *The Borel σ-algebra of $\mathbb{R}$ is defined by*

$$\mathfrak{B}_{\mathbb{R}} := \sigma(\{\text{Open Sets of } \mathbb{R}\})$$

Also we need to know that the generator of $\mathfrak{B}_{\mathbb{R}}$ is not unique. By standard definition, we give

$$\mathfrak{B}_{\mathbb{R}} := \sigma(\{(a,b) : a,b \in \mathbb{R}, a < b\})$$

But we still have the following proposition:

Proposition

Proposition 3.

$$\begin{aligned}
&①\mathfrak{B}_{\mathbb{R}} := \sigma(\{(a,b] : a,b \in \mathbb{R}, a < b\}) \\
&②\mathfrak{B}_{\mathbb{R}} := \sigma(\{[a,b) : a,b \in \mathbb{R}, a < b\}) \\
&③\mathfrak{B}_{\mathbb{R}} := \sigma(\{[a,b] : a,b \in \mathbb{R}, a < b\}) \\
&④\mathfrak{B}_{\mathbb{R}} := \sigma(\{(-\infty,c) : c \in \mathbb{Q}) \\
&⑤\mathfrak{B}_{\mathbb{R}} := \sigma(\{(c,+\infty) : c \in \mathbb{Q})
\end{aligned}$$

I shall illustrate why those are indeed equal. Let's just take a look at ②, we need to show

$$\sigma(\{(a,b) : a,b \in \mathbb{R}, a < b\}) = \sigma(\{[a,b) : a,b \in \mathbb{R}, a < b\})$$

We could indeed prove $L.H.S \subseteq R.H.S$ and $R.H.S \subseteq L.H.S$.

• To show that $L.H.S \subseteq R.H.S$, we need to show

$$(a,b) \in \sigma(\{[a,b) : a,b \in \mathbb{R}, a < b\})$$

and we have

$$(a,b) = \bigcup_{n=1}^{\infty} \left[a + \frac{1}{n}, b\right) \in R.H.S$$

• Similarly,

$$[a,b) = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b\right) \in L.H.S$$

And that's why they are indeed equal.

Definition

Definition 4. *A set G is called a Borel Set if $G \in \mathfrak{B}_{\mathbb{R}}$.*

In fact, any set produced by countable operations are also Borel Sets.

Proposition

Proposition 4. *A set with a single element (or singletons) $\{x\}$ is also a Borel Set.*

Proof. For a singleton $x \in \mathbb{R}$, we have

$$\{x\} = \bigcap_{n=1}^{\infty} \left(x - \frac{1}{n}, x + \frac{1}{n} \right).$$

■

Measures

Given a space X and a σ-algebra $\mathscr{F}$ of subsets of X, $(X,\mathscr{F})$ is called a measurable space.

Definition

Definition 5. *Given a measure space $(X,\mathscr{F})$, define*

$$\mu : \mathscr{F} \longrightarrow [0,+\infty]$$

is a non-negative set function, then μ is a measure if

① $\mu\{\varnothing\} = 0$*;*

② *If $\{A_n, n \geq 1\} \subseteq \mathscr{F}$ such that A_n's are (pairwise) disjoint, then*

$$\mu\left\{\bigcup_{n=1}^{\infty} A_n\right\} = \sum_{n=1}^{\infty} \mu(A_n)$$

this is known as countable additivity.

We say that μ is *finite* if $\mu(X) < +\infty$, we say that μ is a *probability measure* if $\mu(X) = 1$.

We say μ is σ*-finite* if $\exists\{A_n, n \geq 1\} \subseteq \mathscr{F}$ such that $X = \bigcup_{n=1}^{\infty} A_n$ with $\mu(A_n) < +\infty$.

We call the triple $(X,\mathscr{F},\mu)$ a *measure space*.

Here are a few examples of measures on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$:

① A measure μ_1 given by

$$\mu_1 : \mathfrak{B}_{\mathbb{R}} \longrightarrow [0,+\infty], \text{such that } \forall A \in \mathfrak{B}_{\mathbb{R}}, \mu_1(A) = \begin{cases} |A|, \text{if } A \text{ is finite} \\ \\ +\infty, \text{otherwise} \end{cases}$$

We say μ_1 is the *counting measure*.

② Let $x_0 \in \mathbb{R}$, a measure μ_2 is such that

$$\forall A \in \mathfrak{B}_{\mathbb{R}}, \mu_2(A) = \begin{cases} 1, (x_0 \in A) \\ \\ 0, (x_0 \notin A) \end{cases}$$

We say μ_2 is the *probability measure*.

We now give some propositions for the measure space. Below, let $(X, \mathscr{F}, \mu)$ be a measure space.

Proposition

Proposition 5. *(Finite Additivity)*
If $A_1, A_2, \cdots, A_N \in \mathscr{F}$ are disjoint, then

$$\mu\left(A_1 \bigcup A_2 \bigcup \cdots \bigcup A_N\right) = \sum_{n=1}^{N} \mu(A_n).$$

Proposition

Proposition 6. *(Monotoncity)*
Given $A, B \in \mathscr{F}$, if $A \subseteq B$, then $\mu(A) \leq \mu(B)$.

Proof. Note that since $\mu(A), \mu(B)$ could both be infinite, so taking $\mu(B) - \mu(A)$ is incorrect. Indeed, we may construct $B = A \bigcup (B \backslash A)$, where

$$\mu(B) = \mu(A) + \mu(B \backslash A) \geq \mu(A).$$

■

Proposition

Proposition 7. *(Countable / Finite Subadditivity)*
If $\{A_n, n \geq 1\} \subseteq \mathscr{F}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n) \ \ \textit{and} \ \ \mu\left(\bigcup_{n=1}^{N} A_n\right) \leq \sum_{n=1}^{N} \mu(A_n)$$

Note that if A_n's are disjoint, then it should be "$=$".

Proof. Let $B_1 = A_1$, and $B_n := A_n \backslash \left(\bigcup_{i=1}^{n-1} A_i\right)$ for every $n \geq 2$, then we know that $\{B_n, n \geq 1\} \subseteq \mathscr{F}$ and B_n's are disjoint. More importantly, by our construction we have $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. Thus by countable additivity, we have

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(A_n)$$

■

Proposition

Proposition 8. *(Coutinuity From Below)*
Given $\{A_n, n \geq 1\} \subseteq \mathscr{F}$ *such that* $A_n \subseteq A_{n+1}$ *for every n, then*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$

Proof. Set $B_1 = A_1$, $B_n = A_n \backslash A_{n-1}$ for all $n \geq 2$. By definition we have $\{B_n, n \geq 1\} \subseteq \mathscr{F}$ is disjoint, and we have $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$; also for each $N \geq 1$, $\bigcup_{n=1}^{N} B_n = A_n$, which means

$$\begin{aligned}\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) &= \lim_{N\to\infty} \sum_{n=1}^{N} \mu(B_n) \\ &= \lim_{N\to\infty} \mu\left(\bigcup_{n=1}^{N} B_n\right) \\ &= \lim_{n\to\infty} (A_n)\end{aligned}$$

■

Proposition

Proposition 9. *(Continuity From Above)*
Given $\{A_n, n \geq 1\} \in \mathscr{F}$ *such that* $A_{n+1} \subseteq A_n$ *for every* $n \geq 1$*, if* $\mu(A_1) < +\infty$*, then*

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$

Proof. Set $B_n = A_1 \backslash A_n$, then we know that the sequence $\{B_n, n \geq 1\}$ is increasing, then we have

$$\bigcup_{n=1}^{\infty} B_n = A_1 \backslash \bigcap_{n=1}^{\infty} A_n$$

Thus,

$$\mu\left(A_1 \backslash \bigcap_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n) = \lim_{n\to\infty} \mu(A \backslash A_n) \tag{$*$}$$

Given that the measure of A_n is always finite, so $\mu(A_1 \backslash A_n) = \mu(A_1) - \mu(A_n)$ and

$$\mu\left(A_1 \backslash \bigcap_{n=1}^{\infty} A_n\right) = \mu(A_1) - \mu\left(\bigcap_{n=1}^{\infty} A_n\right)$$

Thus, by equation $(*)$,

$$\mu(A_1) - \mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} (\mu(A_1) - \mu(A_n))$$

that is

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$

■

Construction of Lebesgue Measure on $\mathbb{R}$

Now, we want a measure on the real line $\mathbb{R}$ such that:

① It assigns to an interval the length of that interval;

② It is translation invariant (i.e $A, A+x$ has the same measure).

We begin with the definition of an *outer measure*:

Definition

Definition 6. $\forall A \subseteq \mathbb{R}$, *define a set function*

$$m^*(A) = \inf\left\{\sum_{n=1}^{\infty} l(I_n) : A \subseteq \bigcup_{n=1}^{\infty} I_n\right\}$$

where I_n's are open intervals. Then m^* is called an outer measure on $\mathbb{R}$ (The Lebesgue Outer Measure), $l(I_n)$ represents the length of that interval.

Based on the definition, we have the following properties:

Proposition

Proposition 10. ① $\forall A \subseteq \mathbb{R}$, $m^*(A) \geq 0$ *and we define* $m^*(\varnothing) = 0$*;*

② *(Monotoncity) If $A, B \subseteq \mathbb{R}$ and $A \subseteq B$, then $m^*(A) \leq m^*(B)$;*

③ *(Countable Subadditivity) If $\{A_n, n \geq 1\}$ is a sequence of subsets of $\mathbb{R}$, then*

$$m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} m^*(A_n)$$

Proof. If $m^*(A) = +\infty$, then ③ holds trivially, therefore we may assume that $m^*(A_n)$ is finite for every n.

For each $n \geq 1$, $\forall \varepsilon > 0$, one can choose a sequence of open intervals $\{I_{ni}, i \geq 1\}$ such that

$$A_n \subseteq \bigcup_{i=1}^{\infty} I_{ni} \text{ and } \sum_{i=1}^{\infty} l(I_{ni}) \leq m^*(A_n) + \frac{\varepsilon}{2^n}$$

therefore

$$\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} \bigcup_{i=1}^{\infty} I_{ni}$$

thus by definition

$$m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n,i} l(I_{ni}) = \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} l(I_{ni})$$

since

$$\sum_{n=1}^{\infty} \sum_{i=1}^{\infty} l(I_{ni}) \leq \sum_{n=1}^{\infty} \left(m^*(A_n) + \frac{\varepsilon}{2^n}\right) = \sum_{n=1}^{\infty} m^*(A_n) + \varepsilon$$

which is

$$m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} m^*(A_n) + \varepsilon$$

Knowing that ε is arbitrarily small, thus

$$m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} m^*(A_n)$$

■

Remark: In fact, any set function on $2^{\mathbb{R}}$ that satisfies properties ①,②,③ is called an **outer measure**.

More properties follow:

Proposition

Proposition 11. *If $I \subseteq \mathbb{R}$ is an interval, then $m^*(I) = l(I)$;*

Proof. First, we consider $I = [a,b]$ for some $a, b \in \mathbb{R}, a < b$. Then, $\forall \varepsilon > 0$, we construct

$$I_1 = (a - \varepsilon, b + \varepsilon)$$

clearly $I \subseteq I_1$, thus

$$m^*(I) \leq m^*(I_1) \leq l(I_1) = (b-a) + 2\varepsilon$$

Since ε is arbitrarily small, so

$$m^*(I) \leq (b-a) = l(I)$$

Now, we need to show that $m^*(I) \geq l(I)$, which is sufficient to show that $\forall \{I_n, n \geq 1\}$ open and covering I, we have

$$\sum_{n=1}^{\infty} l(I_n) \geq l(I)$$

Let $\{I_n, n \geq 1\}$ be such a covering, since $I = [a,b]$ is compact, thus I can be covered by finitely many I_n's. We can extract a finite collection $I_1, I_2, \cdots, I_N$ with $I_n = (a_n, b_n)$, such that $a_1 < a; b_N > b$ and $a_n < b_{n-1}$ for every $n \geq 2$. Then

$$\sum_{n=1}^{\infty} l(I_n) \geq \sum_{i=1}^{N} l(I_n) = b_1 - a_1 + \sum_{n=2}^{N}(b_n - a_n) \geq b_1 - a_1 + \sum_{n=2}^{N}(b_n - b_{n-1}) = b_N - a_1 \geq b - a = l(I)$$

Now assume I is a finite integral with end points $a, b, a < b$. Let $\varepsilon > 0$, then

$$[a+\varepsilon, b-\varepsilon] \subseteq [a-\varepsilon, b+\varepsilon], \left(\varepsilon < \frac{b-a}{2}\right)$$

By monotoncity, we have

$$b - a - 2\varepsilon \leq m^*(I) \leq b - a + 2\varepsilon$$

thus

$$m^*(I) = b - a = l(I)$$

Finally, consider I to be an infinite integral, thus $\forall M > 0$, there exists a closed finite interval I_M such that $I_M \subseteq I$ and $l(M) \leq M$, which means

$$m^*(I) \geq m^*(I_M) \leq M$$

since M is arbitrarily large, thus $m^*(I) = +\infty$. ■

Proposition

Proposition 12. *m^* is translation invariant, i.e $\forall A \subseteq \mathbb{R}, x \in \mathbb{R}$, $m^*(A) = m^*(A+x)$;*

Proposition

Proposition 13. *$\forall A \subseteq \mathbb{R}$, $m^*(A) := \inf\{m^*(B), B \subseteq \mathbb{R}, A \subseteq B\}$ where B is open.*

Proof. $\forall B$ open with $A \subseteq B$, by monotoncity we have $m^*(A) \leq m^*(B)$. *W.L.O.G*, assume $m^*(A) < +\infty$, that means $\forall \varepsilon > 0$, there exists open intervals $\{I_n, n \geq 1\}$ such that

$$A \subseteq \bigcup_{n=1}^{\infty} I_n \text{ and } m^*(A) + \varepsilon \geq \sum_{n=1}^{\infty} l(I_n)$$

Now set $B = \bigcup_{n=1}^{\infty} I_n$ thus we have

$$m^*(B) = m^*\left(\bigcup_{n=1}^{\infty} I_n\right) \leq \sum_{n=1}^{\infty} m^*(I_n) = \sum_{n=1}^{\infty} l(I_n) \leq m^*(A) + \varepsilon$$

since ε is arbitrarily small, thus we have $m^*(A) := \inf\{m^*(B), B \subseteq \mathbb{R}, A \subseteq B\}$ where B is open.

■

Proposition

Proposition 14. *If $A = A_1 \bigcup A_2 \subseteq \mathbb{R}$ with $d(A_1, A_2) > 0$, then $m^*(A) = m^*(A_1) + m^*(A_2)$.*

Proof. By definition, it holds trivially that $m^*(A) \leq m^*(A_1) + m^*(A_2)$, we remain to show that $m^*(A) \geq m^*(A_1) + m^*(A_2)$. We may also assume that $m^*(A) \leq \infty$.

$\forall \varepsilon > 0$, $\exists$ a sequence of open intervals I_n such that $A \subseteq \bigcup_{n=1}^{\infty} I_n$ and

$$\sum_{n=1}^{\infty} l(I_n) \leq m^*(A) + \varepsilon$$

also, it is possible to further rewrite each I_n as $I_n \subseteq \bigcup_{i=1}^{\infty} I_{ni}$ where I_{ni} are open, $n \geq 1$ and $l(I_{ni}) < d(A_1, A_2)$, also

$$\sum_{i=1}^{\infty} l(I_{ni}) \leq l(I_n) + \varepsilon$$

Now rename I_{ni} as J_m, since $l(J_m) < d(A_1, A_2)$, which means J_m can intersect at most one of A_1, A_2, then define $p = \{1, 2\}$, set

$$M_p = \{m : J_m \bigcap A_p \neq \varnothing\}$$

then $\{J_m : m \in M_p\}$ is an open interval covering of A_p. We now know that

$$\begin{aligned} m^*(A_1) + m^*(A_2) \leq \sum_{m \in M_1} l(J_m) + \sum_{m \in M_2} l(J_m) &\leq \sum_{m=1}^{\infty} J_m \\ &= \sum_n \sum_i l(I_{ni}) \\ &\leq \sum_n \left(l(I_n) + \frac{\varepsilon}{2^n} \right) \\ &= m^*(A) + 2\varepsilon \end{aligned}$$

■

Proposition

Proposition 15. *If $A = \bigcup_{r=1}^{\infty} J_r$ where J_r's are almost disjoint intervals, i.e they at most intersect at boundary points. Then*

$$m^*(A) = \sum_{k=1}^{\infty} l(J_k)$$

Proof. If $l(J_k) = \infty$ for some $k \geq 1$, then since $J_k \subseteq A$, it implies that $m^*(A) = \infty$, in this case the desired relation holds trivially. Thus we may assume that $l(J_k) < \infty$ for all $k \geq 1$. Fix an arbitrary $\varepsilon > 0$, for all $k \geq 1$, choose an open interval $I_k \subseteq J_k$ such that

$$l(J_k) \leq l(I_k) + \frac{\varepsilon}{2^k}$$

Furthermore, for any fixed integer n, we can make the choice of $I_1, I_2, \cdots, I_n$ such that they are disjoint with positive distance from one another, thus

$$\bigcup_{k=1}^{n} I_k \subseteq \bigcup_{k=1}^{n} J_n \subseteq A$$

and

$$m^*(A) \geq m^*\left(\bigcup_{k=1}^{n} I_k\right) = \sum_{k=1}^{n} l(I_k) \geq \sum_{k=1}^{n}\left(l(J_k) - \frac{\varepsilon}{2^k}\right) \geq \sum_{k=1}^{n} l(J_k) - \varepsilon$$

Since ε is arbitrarily small, n is arbitrarily large, in fact

$$m^*(A) \geq \sum_{k=1}^{\infty} l(J_k)$$

Also by definition, $m^*(A) \leq \sum_{k=1}^{\infty} l(J_k)$ holds trivially. ■

Defining Measurable Sets

Definition

Definition 7. *A set $M \subseteq \mathbb{R}$ is called m^*-measurable, if $\forall B \subseteq \mathbb{R}$,*

$$m^*(B) = m^*(B \cap M) + m^*(B \cap M^C)$$

Otherwise, M is called non-measurable.

Remarks :

① By subadditivity, we already know that $\forall A, B \subseteq \mathbb{R}$,

$$m^*(B) \leq m^*(B \cap A) + m^*(B \cap A^C)$$

thus to show that A is m^*-measurable, we only need to show that

$$m^*(B) \geq m^*(B \cap A) + m^*(B \cap A^C)$$

② Whether or not A is m^*-measurable depends on whether or not $<$ can occur for some B.

③ Non-measurable set exists, it relies on the *Axioms of Choice* (AC).

Theorem

Theorem 2. *(Carathéodory's)*
Let $\mathscr{M} := \{A \subseteq \mathbb{R}\}$ where A is m^*-measurable, then $\mathscr{M}$ is a σ-algebra of subsets of $\mathbb{R}$, define

$$m : \mathscr{M} \longrightarrow [0, +\infty], \forall A \in \mathscr{M}, m(A) = m^*(A)$$

then m is a measure called the Lebesgue Measure on $\mathbb{R}$. If $A \in \mathscr{M}$, we say A is (Lebesgue) measurable.

Proof. First we show that $\mathscr{M}$ is a σ-algebra. The first two properties are trivial, thus we remain to show that if $\{A_n : n \geq 1\} \subseteq \mathscr{M}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathscr{M}$.

To prove this, we first check that $\mathscr{M}$ is closed under finite union, i.e if $A_1, A_2, \cdots, A_N \in \mathscr{M}$, then $\bigcup_{n=1}^{N} A_n \in \mathscr{M}$.

It is sufficient to show $A_1 \bigcup A_2 \in \mathscr{M}$ since we may perform the induction on N. By definition,

$$\begin{aligned}\forall B \in \mathbb{R}, m^*(B) &= m^*(B \cap A_1) + m^*(B \cap A_1^C) \\ &= m^*(B \cap A_1) + m^*((B \cap A_1^C) \cap A_2) + m^*((B \cap A_1^C) \cap A_2^C)\end{aligned}$$

Notice that $(B \cap A_1) \cup (B \cap A_1^C \cap A_2) = B \cap (A_1 \cup A_2)$, then

$$m^*(B) \geq m^*(B \cap (A_1 \cup A_2)) + m^*(B \cap (A_1 \cup A_2)^C)$$

Now consider the countable sequence $\{A_n : n \geq 1\} \subseteq \mathscr{M}$, set

$$B_1 = A_1, B_n = A_n \Big\backslash \left(\bigcup_{i=1}^{n-1} A_i \right), n \geq 2$$

then we notice that B_n's are disjoint, and $\bigcup_{i=1}^{\infty} B_n = \bigcup_{i=1}^{\infty} A_n$ where $B_n \in \mathscr{M}, n \geq 1$. To show that $\bigcup_{n=1}^{\infty} A_n \in \mathscr{M}$, we only need to show that $\bigcup_{n=1}^{\infty} B_n \in \mathscr{M}$. For every $n \geq 1$, let $E_n = \bigcup_{i=1}^{n} B_i$, thus we know that $E_n \in \mathscr{M}, n \geq 1$.

Now $\forall B \in \mathbb{R}, n \geq 1$, we have

$$m^*(B) = m^*(B \cap E_n) + m^*(B \cap E_n^C)$$

where $E_n \subseteq \bigcup_{n=1}^{\infty} B_n$ and $\left(\bigcup_{n=1}^{\infty} B_n \right)^C \subseteq E_n^C$, that is

$$\begin{aligned}m^*(B) &\geq m^*(B \cap E_n \cap B_n) + m^*(B \cap E_n \cap B_n^C) + m^*(B \cap \left(\bigcup_{n=1}^{\infty} B_n \right)^C) \\ &\geq m^*(B \cap B_n) + m^*(B \cap E_{n-1}) + m^*(B \cap \left(\bigcup_{n=1}^{\infty} B_n \right)^C) \\ &\vdots \\ &\geq \sum_{i=1}^{n} m^*(B \cap B_i) + m^*(B \cap \left(\bigcup_{n=1}^{\infty} B_n \right)^C)\end{aligned}$$

As $n \to \infty$, we have

$$\begin{aligned}m^*(B) &\geq \sum_{n=1}^{\infty} m^*(B \cap B_n) + m^*(B \cap \left(\bigcup_{n=1}^{\infty} B_n \right)^C) \\ &\geq m^*(B \cap \left(\bigcup_{n=1}^{\infty} B_n \right)) + m^*(B \cap \left(\bigcup_{n=1}^{\infty} B_n \right)^C)\end{aligned}$$

thus $\mathscr{M}$ is a σ-algebra.

Next, we will check that $m = m^*|_{\mathscr{M}}$ is a measure.

Since $m(A) \geq 0$ and $m(\varnothing) = 0$ are obvious, we will only check the countable additivity.

Given $\{A_n, n \geq 1\} \subseteq \mathscr{M}$ where A_i's are disjoint, we follow the ame argument as above, and we get for every $n \geq 1$,

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

Therefore, for every $m\left(\bigcup_{i=1}^{\infty} A_i\right) \geq m\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} m(A_i)$, as $n \longrightarrow \infty$,

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

and by subadditivity, the other direction is trivial, which means $\mathscr{M}$ is a measure. ■

Proposition

Proposition 16. *$\mathscr{M}$ and m are translation invariant, i.e $\forall A \in \mathscr{M}, x \in \mathbb{R}, A + x \in \mathscr{M}$ and $m(A) = m(A+x)$.*

Proof.

$$\begin{aligned}\forall B \in \mathbb{R}, m^*(B) &= m^*(B-x)\\ &= m^*(B - x \cap A) + m^*(B - x \cap A^C)\\ &= m^*(B \cap A + x) + m^*(B \cap (A+x)^C)\end{aligned}$$

■

Theorem

Theorem 3. *$\forall a, b \in \mathbb{R}, a < b$, then $(a,b) \in \mathscr{M}$ and*

$$m((a,b)) = b - a$$

Corollary

Corollary 1. *Every Borel Set is Lebesgue Measurable.*

Properties of a Lebesgue Measure

Now we shall discuss some important properties for a Lebesgue measure m.

Proposition

Proposition 17. *(Translation Invariant)*
$\forall A \in \mathscr{M}, x \in \mathbb{R}$, then $A + x \in \mathscr{M}$ and $m(A) = m(A + x)$.

Proof. We have proved this in the previous lecture. ■

Proposition

Proposition 18. *(Any Interval Is Measurable)*
$\forall a, b \in \mathbb{R}, a < b$, we have $(a,b) \in \mathscr{M}$ and $m((a,b)) = b - a$.

Proof. Let $a < b, a, b \in \mathbb{R}$, then $\forall B \subseteq \mathbb{R}$ we only need to show

$$m^*(B) = m^*(B \cap (a,b)) + m^*(B \cap (a,b)^C)$$

As mentioned previously, we only need to prove

$$m^*(B) \geq m^*(B \cap (a,b)) + m^*(B \cap (a,b)^C)$$

Assume that $m^*(B) < +\infty$, then $\forall \varepsilon > 0$, one can always find a sequence of open intervals $\{I_n, n \geq 1\}$ such that

$$B \subseteq \bigcup_{n=1}^{\infty} I_n \text{ and } m^*(B) + \varepsilon \geq \sum_{n=1}^{\infty} l(I_n)$$

For each $n \geq 1$, let $J_n = I_n \cap (a,b)$; $K_n = I_n \cap (-\infty, a]$; $L_n = I_n \cap [b, +\infty)$, and for any fixed n, J_n, K_n, L_n are almost disjoint, hence

$$m^*(I_n) = m^*(J_n) + m^*(K_n) + m^*(L_n)$$

so we have

$$\begin{aligned}
m^*(B) \geq \sum_{n=1}^{\infty} l(I_n) - \varepsilon &= \sum_{n=1}^{\infty} l(J_n) + \sum_{n=1}^{\infty} (l(K_n) + l(L_n)) - \varepsilon \\
&= \sum_{n=1}^{\infty} m^*\left(I_n \bigcap (a,b)\right) + \sum_{n=1}^{\infty} m^*\left(I_n \bigcap (a,b)^C\right) - \varepsilon \\
&\geq m^*\left(\bigcup_{n=1}^{\infty} I_n \bigcap (a,b)\right) + m^*\left(\bigcup_{n=1}^{\infty} I_n \bigcap (a,b)^C\right) - \varepsilon \\
&\geq m^*(B \cap (a,b)) + m^*(B \cap (a,b)^C) - \varepsilon
\end{aligned}$$

where ε is chosen arbitrarily small. ■

We have the following regularity properties for Lebesgue measure m:

Proposition

Proposition 19. $\forall A \in \mathscr{M}, \varepsilon > 0$, *there exists an open set* G *where* $A \subseteq G$ *such that*

$$m(G\backslash A) < \varepsilon$$

Proof. It is equivalent to prove that $m(G) - m(A) < \varepsilon$. Since we also have $m(G) = m(A) + m(G\backslash A)$, so when the measure is not infinity, we are done. Now consider when the measure is infinity, that is assume $m(A) = +\infty$. Then consider the sequences $A_n := A \bigcap [-n,n]$, where $m(A_n) < +\infty$ and there exists an open set such that $A_n \subseteq G_n$ where $m(G_n\backslash A_n) < \varepsilon/2^n$. The let $G = \bigcup_{n=1}^{\infty} G_n$, $A = \bigcup_{n=1}^{\infty} A_n$, we now have G is open and $A \subseteq G$, also

$$m(G\backslash A) = m\left(\bigcup_{n=1}^{\infty}(G_n\backslash A) \leq \sum_{n=1}^{\infty} m(G_n\backslash A)\right) \leq \sum_{n=1}^{\infty} m(G_n\backslash A_n) < \varepsilon$$

■

Proposition

Proposition 20. $\forall A \in \mathscr{M}, \varepsilon > 0$, *there exists a closed set* F *where* $F \subseteq A$ *such that*

$$m(A\backslash F) < \varepsilon$$

Proof. to be implemented. ■

Proposition

Proposition 21. *(Outer Regular)*
$\forall A \in \mathscr{M}$, *we have*

$$m(A) := \inf\{m(G) : G \text{ is open}, A \subseteq G\}$$

Proof. to be implemented. ■

Proposition

Proposition 22. *(Inner Regular)*
$\forall A \in \mathscr{M}$, *we have*

$$m(A) := \sup\{m(K) : K \text{ is compact}, K \subseteq A\}$$

Proof. to be implemented ■

Proposition

Proposition 23. $\forall A \in \mathscr{M}$ *and* $m(A) < \infty$, *then* $\forall \varepsilon > 0$ *there exists a compact set* $K \subseteq A$ *such that*

$$m(A\backslash K) < \varepsilon$$

Proof. By (ii), we know that $\forall \varepsilon > 0$, there exists a closed set $F \subseteq A$ such that

$$m(A \backslash F) < \frac{\varepsilon}{2}$$

Now for each n, set $F_n = [-n,n] \cap F$, now we know that F_n is compact, hence $A \backslash F_n$ is decreasing and $\bigcap_{n=1}^{\infty}(A \backslash F_n) = A \backslash F$, by continuity from above,

$$m(A \backslash F) = \lim_{n\to\infty} m(A \backslash F_n) \leq \frac{\varepsilon}{2}$$

then there exists a number N such that $m(A \backslash F_N) < \varepsilon$, since F_n is compact, $F_n \subseteq A$, thus

$$m(A \backslash F_N) < \varepsilon$$

■

Proposition

Proposition 24. *$\forall A \in \mathscr{M}$ and $m(A) < +\infty$, then $\forall \varepsilon > 0$, there exists a finite collection of intervals $I_1, I_2, \cdots, I_N$ such that*

$$m\left(A \lhd \left(\bigcup_{n=1}^{N} I_n\right)\right) < \varepsilon$$

we define $A \lhd B = (A \backslash B) \bigcap (B \backslash A)$.

Proof. $\forall \varepsilon > 0$, choose $\{I_n, n \geq 1\}$ to be a sequence of open intervals such that $A \subseteq \bigcup_{n=1}^{\infty} I_n$ and

$$m(A) > \sum_{n=1}^{\infty} l(I_n) - \frac{\varepsilon}{2}$$

since $m(A) < +\infty$, so the sequence $\sum\limits_{n=1}^{\infty} l(I_n)$ is convergent so we have N large enough such that

$$\sum_{n=N+1}^{\infty} l(I_n) < \frac{\varepsilon}{2}$$

then we have

$$\begin{aligned}
m\left(A \lhd \bigcup_{n=1}^{N} I_n\right) &= m\left(A \Big\backslash \bigcup_{n=1}^{N} I_n\right) + m\left(\bigcup_{n=1}^{N} I_n \Big\backslash A\right) \\
&\leq m\left(\bigcup_{n=1}^{\infty} I_n \Big\backslash \bigcup_{n=1}^{N} I_n\right) + m\left(\bigcup_{n=1}^{\infty} I_n \Big\backslash A\right) \\
&= m\left(\bigcup_{n=N+1}^{\infty} I_n\right) + m\left(\bigcup_{n=1}^{\infty} I_n \Big\backslash A\right) \\
&\leq \frac{\varepsilon}{2} + \sum_{n=N+1}^{\infty} l(I_n) - m(A) = \varepsilon
\end{aligned}$$

■

Proposition

Proposition 25. *(Completeness)*
Given $(\mathbb{R},\mathscr{M},m)$ is a complete measure space in the sense that $\forall A\subseteq\mathbb{R}$ if $\forall B\in\mathscr{M}$ such that $A\subseteq B$ and $m(B)=0$, then $A\in\mathscr{M}$ and $m(A)=0$. i.e any subset of a null set is again a null set.

Proof. Assume $A\subseteq\mathbb{R}$ such that $A\subseteq B$ for some $B\in\mathscr{M}$ with $m(B)=0$. We want to show that $\forall E\subseteq\mathbb{R}$,

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

which means we only need to show

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

Since $A\subseteq B$, we then have $m^*(E\cap A)\leq m^*(E\cap B)$, since $B\in\mathscr{M}$, thus

$$\begin{aligned}m^*(E\cap A^C)&=m^*(E\cap A^C\cap B)+m^*(E\cap A^C\cap B^C)\\&\leq m^*(B)+m^*(E\cap B^C)\end{aligned}$$

then

$$m^*(E\cap A)+m^*(E\cap A^C)\leq m^*(E\cap B)+m^*(E\cap B^C)=m^*(E)$$

thus $\mathscr{A}\in\mathscr{M}$ and it follows that $m(A)=0$. ■

Remark: An equivalent statement of saying $(\mathbb{R},\mathscr{M},m)$ is complete is: $\forall E\subseteq\mathbb{R}$, if $F,G\in\mathscr{M}$ such that $F\subseteq E\subseteq G$ and $m(G\backslash F)=0$, then $E\in\mathscr{M}$ and $m(E)=m(G)=m(F)$.

Uniqueness of m

Let's consider $m\Big|_{\mathfrak{B}_{\mathbb{R}}}$ where $(\mathbb{R},\mathfrak{B}_{\mathbb{R}},m)$ is a measure space, we have the following proposition:

Proposition

Proposition 26. *Up to re-scaling, the Lebesgue measure m is the unique, non-trivial measure on $(\mathbb{R},\mathfrak{B}_{\mathbb{R}})$ such that it's finite on compact sets and translation invariant.*

That is, if μ is another measure on $(\mathbb{R},\mathfrak{R}_{\mathbb{R}})$, then $\mu=cm$ for $c\in\mathbb{R}$.

The uniqueness of m can be derived from the famous *Dynkin's $\pi-d$ Theorem*, which we will introduce later.

Theorem

Theorem 4. *Given a space X, let $\mathscr{C}$ be a collection of the subset of $\mathscr{C}$, then $\mathscr{C}$ is called a π-system if $A,B\in\mathscr{C}$ implies $A\bigcap B\in\mathscr{C}$. Now assume $\mathscr{F}=\sigma(\mathscr{C})$, a σ-algebra generated by this π-system and μ_1,μ_2 are 2 finite measures on $(X,\mathscr{F})$ such that $\mu_1(X)=\mu_2(X)$ and $\mu_1=\mu_2$ on $\mathscr{C}$, then $\mu_1=\mu_2$ for the entire $\mathscr{F}$.*

Recall that $\mathfrak{B}_{\mathbb{R}} := \sigma(\{(a,b) : a,b \in \mathbb{R}, a < b\})$, we claim that $\mathfrak{B}_{\mathbb{R}} \bigcap \{\varnothing\}$ is a π-system.

For each $n \geq 1$, consider the set $\mathscr{C}_{\cap[-n,n]} := \{(a,b) \bigcap [-n,n] : a,b \in \mathbb{R}\}$, it is again a π-*system* and we have that

$$\sigma(\mathscr{C}_{\cap[-n,n]}) = \sigma(\mathscr{C})_{\cap[-n,n]}$$

which implies

$$\mathfrak{B}_{[-n,n]} = \sigma(\mathscr{C}_{\cap[-n,n]})$$

Now we give the formal definition of a π-system and a d-system.

Definition

Definition 8. *Let X to be a space (non-empty sets), let $\mathscr{I}$ be a collection of subsets of X, then $\mathscr{I}$ is a π-system if $\forall A, B \in \mathscr{I}$, $A \bigcap B \in \mathscr{I}$*

Definition

Definition 9. *Let X to be a space (non-empty sets), let $\mathscr{D}$ be a collection of subsets of X, then $\mathscr{D}$ is a d-system if the followings hold:*

① *$X \in \mathscr{D}$;*

② *If $A, B \in \mathscr{D}$ with $A \subseteq B$, then $B \backslash A \in \mathscr{D}$;*

③ *If $\{A_n : n \geq 1\} \in \mathscr{D}$ and the sequence is increasing, i.e $A_n \subseteq A_{n+1}$, then $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{D}$.*

Theorem

Theorem 5. *Let X be a space and $\mathscr{F}$ be a collection of subsets of X, then $\mathscr{F}$ is a σ-algebra if and only if $\mathscr{F}$ is a π-system and $\mathscr{F}$ is a d-system.*

Proposition

Proposition 27. *If μ is a measure on $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}})$ such that $\forall I \in \mathbb{R}$, we have $\mu(I) = l(I)$, then μ is the Lebesgue measure m.*

Proof. $\forall n \geq 1$, consider μ restricted on $\mathfrak{B}_{[-n,n]}$, clearly

$$\mu([-n,n]) = m([-n,n]) = 2n$$

we have that $\mathscr{C}_{\cap[-n,n]}$ as the generating π-system of $\mathfrak{B}_{\mathbb{R}}$, Now $\forall a, b \in \mathbb{R}$, since $\mu(\varnothing) = m(\varnothing) = 0$, then

$$\mu((a,b) \cap [-n,n]) = l((a,b) \cap [-n,n]) = m((a,b) \cap [-n,n])$$

according to the theorem, $\mu = m$ on $\mathfrak{B}_{[-n,n]}$, then

$$\mu(A) = \lim_{n \to \infty} \mu(A_n) = \lim_{n \to \infty} m(A_n) = m(A_n)$$

■

To prove the proposition ⑤ (Uniqueness of m), we first need to prove that $\mathfrak{B}_{\mathbb{R}}$ is translation invariant.

We may use the "Good Set Strategy", that is we first define some "good set", and then we try to show that everything we need to prove is indeed inside the "good set".

We define the collection of good set as:

$$\forall x \in \mathbb{R}, \Sigma := \{B \in \mathfrak{B}_{\mathbb{R}} : B + x \in \mathfrak{B}_{\mathbb{R}}\}$$

the we show that Σ is a σ-algebra, once Σ is a σ-algebra, we already know that

$$\{(a,b) : a,b \in \mathbb{R}\} \subseteq \Sigma \subseteq \mathfrak{B}_{\mathbb{R}}$$

then we can easily show that

$$\mathfrak{B}_{\mathbb{R}} = \sigma(\{(a,b) : a,b \in \mathbb{R}, a < b\}) \subseteq \Sigma$$

then we can show that $\Sigma = \mathfrak{B}_{\mathbb{R}}$.

Now, we prove the uniqueness of m:

Proof. Set $c = \mu([0,1]), c > 0$. Since if $c = 0$ then μ is trivial.

$$\forall n \geq 1, \mu\left(\left(0, \frac{1}{n}\right]\right) = \frac{c}{n}$$

$$\forall m = 1, 2, \cdots, n-1, \mu\left(\left(0, \frac{m}{n}\right]\right) = \frac{m}{n}c$$

$$\forall q \in \mathbb{Q}, \cap (0,1], \mu((0,q]) = qc$$

$$\forall q \in \mathbb{Q}^{+}, \mu((0,q]) = qc$$

$$\forall a \in \mathbb{R}, \mu((a, a+q]) = qc$$

since $\mathbb{Q}$ is dense in $\mathbb{R}$, thus for any interval $I \subseteq \mathbb{R}$, $\mu(I) = cl(I)$ by continuity.

Also, $\forall, n \geq 1, ab \in \mathbb{R}, a < b$, we have

$$\mu((a,b) \cap [-n,n]) = cl((a,b) \cap [-n,n]) = cm((a,b) \cap [-n,n])$$

thus we have $\mu = cm$ on $\mathfrak{B}_{\mathbb{R}}$, and we can conclude that $\mu = cm$ on $\mathfrak{B}_{\mathbb{R}}$. ■

Scaling Property

Proposition

Proposition 28. *The Lebesgue measure m has the scaling property that $\forall A \in \mathscr{M}, c \in \mathbb{R}$, we have*

$$cA := \{cx, x \in A\} \in \mathscr{M}, m(cA) = |c| m(A)$$

First proof : (Direct Proof)

Proof. Key points here is : Given $A \subseteq \mathbb{R}, \{I_n, n \geq 1\}$ is an open interval covering of A if and only if $\{cI_n, n \geq 1\}(c \neq 0)$ is an open interval covering of cA and $l(cI_n) = |c|l(I_n)$ also $m^*(cA) = |c|m^*(A)$ ■

Second proof:

Proof. Alternatively, we can prove the scaling property on $\mathfrak{B}_{\mathbb{R}}$, $\forall B \in \mathfrak{B}_{\mathbb{R}}$, define

$$\mu_c(B) = m(cB)$$

and we can easily check that μ_c is indeed a measure and it's translation invariant, and $\mu_c = |c|m$ on $\mathfrak{B}_{\mathbb{R}}$, it means

$$\forall B \in \mathfrak{B}_{\mathbb{R}}, c \in \mathbb{R}\backslash 0, m(cB) = |c|m(B)$$

and we have something different...

How can we go from $\mathfrak{B}_{\mathbb{R}}$ to $\mathscr{M}$?

Relation Between $\mathfrak{B}_{\mathbb{R}}$ and $\mathscr{M}$

Recall that $(\mathbb{R}, \mathscr{M}, m)$ is complete, i.e $\forall B \in \mathbb{R}$, if $\exists A \in \mathscr{M}$ such that $m(A) = 0$. If $B \subseteq A$ then $B \in \mathscr{M}$ and $m(B) = 0$.

Definition

Definition 10. *Given a measure space $(X, \mathscr{F}, \mu)$, consider the following collection of subsets of X, called $\mathscr{N}$, such that*

$$\mathscr{N} := \{B \subseteq X : \exists A \in \mathscr{F}, \mu(A) = 0, B \subseteq A$$

then define $\overline{\mathscr{F}} := \sigma(\mathscr{F} \bigcup \mathscr{N})$, called the completion of $\mathscr{F}$ with respect to μ, if $\mathscr{N} \subseteq \mathscr{F}$, then $\overline{\mathscr{F}} = \mathscr{F}$ in which case $\mathscr{F}$ is already complete.

Proposition

Proposition 29. *We have*

$$\overline{\mathscr{F}} := \{F \subseteq X; \exists E, G \in \mathscr{F}, E \subseteq F \subseteq G, \mu(G\backslash E) = 0\}$$

Proof. Denote $\mathscr{R}$ to be the collection on the $R.H.S$, we first check that $\mathscr{R}$ is a σ-algebra. (to be implemented) ■

Definition

Definition 11. *Given a measure space $(X, \mathscr{F}, \mu)$, then it can be extended to $\overline{\mathscr{F}}$ as follows:*

If $F \in \overline{\mathscr{F}}$, assume $E, G \in \mathscr{F}$ such that $E \subseteq F \subseteq G$ where $\mu(G \backslash E) = 0$, then we define

$$\mu(E) = \mu(F) = \mu(G)$$

and now $(X, \overline{\mathscr{F}}, \mu)$ is a complete measure space.

Theorem

Theorem 6. *$(\mathbb{R}, \mathscr{M}, m)$ is the completion of $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}}, m)$.*

Proof. Given $A \in \mathscr{M}$, then $\forall n \geq 1$, there exists an open set G_n such that $A \subseteq G_n$ and

$$m^*(G_n \backslash A) < \varepsilon = \frac{1}{n}$$

Similarly, there also exists a closed set F_n such that $F_n \subseteq A$ and

$$m^*(A \backslash F) < \varepsilon = \frac{1}{n}$$

Set

$$C := \bigcap_{n=1}^{\infty} G_n, B := \bigcup_{n=1}^{\infty} F_n$$

then by definition, $B, C \in \mathfrak{B}_{\mathbb{R}}$, we know that $\forall n \geq 1$,

$$m(C \backslash A) \leq m(G_n \backslash A) < \frac{1}{n}$$

$$m(A \backslash B) \leq m(A \backslash F_n) < \frac{1}{n}$$

meaning that

$$m(C \backslash B) = m(C) - m(B) < \frac{2}{n}$$

and the choice of n is arbitrarily large.

Thus $m(C \backslash B) = 0$ as $n \longrightarrow \infty$, meaning that $A \in \mathfrak{B}_{\mathbb{R}}$, and $\mathscr{M} \in \overline{\mathfrak{B}_{\mathbb{R}}}$ and $\mathscr{M} = \overline{\mathfrak{B}_{\mathbb{R}}}$. ■

Corollary

Corollary 2. *Any measurable set is different from a Borel set by at most a null set, meaning that if $m = (cB) = |c|m(B)$ for all $B \in \mathfrak{B}_{\mathbb{R}}$, then $m(cA) = |c|m(A)$ for all $A \in \mathscr{M}$.*

Some Special Sets

Now we have the following question:

Is there a subset $A \subseteq \mathbb{R}$ that is measurable with $m(A) = 0$, such that A is uncountable? (*It is a fact that if A is countable, then $m(A) = 0$*).

The answer is yes, *Cantor Set* is such an example.

1.8.1 An uncountable set with measure 0 : Cantor Set

The Cantor Set is given by

$$C_0 = [0,1]$$

$$C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$$

$$C_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]$$

$$\vdots$$

$$\vdots$$

It is generated by removing the middle thirds each time. It is obvious that C_n is decreasing. Let

$$\mathscr{C} := \bigcap_{n=0}^{\infty} C_n$$

to be our Cantor Set, then we claim the following facts:

① Cantor Set is closed;

② The measure of the Cantor Set is 0, since we see that

$$m(C_n) = 2^n \frac{1}{3^n} = \left(\frac{2}{3}\right)^n$$

then obviously

$$m(\mathscr{C}) = \lim_{n \to \infty} m(C_n) = 0$$

Also, observe that $\forall x \in [0,1]$, we can find a sequence $(a_1, a_2, \cdots)$ where $a_i \in \{0,1,2\}$ such that

$$x = \sum_{n=1}^{\infty} a_n \frac{1}{3^n}$$

We assign the value 0, if it is in the first thirds, 1 for the second thirds and 2 for the third thirds, and we notice that some numbers may have more than 1 expansion, such as

$$\frac{1}{3} = \begin{cases} (1,0,0,0,\cdots) \\ (0,2,2,2,\cdots) \end{cases}$$

and

$$\frac{2}{3} = \begin{cases} (1,2,2,2,\cdots) \\ (2,0,0,0,\cdots) \end{cases}$$

Thus we may define the Cantor Set as

$$\mathscr{C} := \{x \in [0,1], \text{ where the base expansion of } x \text{ does not contain the digit } 1\}$$

numbers like $\frac{1}{3}$ is still in $\mathscr{C}$ because it has a different expansion that does not contain 1, then

$$\mathbf{Card}(\mathscr{C}) = |\{a_n, n \geq 1\} : a_n \in \{0,2\}| = \mathbf{Card}([0,1])$$

meaning that $\mathscr{C}$ is uncountable.

Define $f : \mathscr{C} \longrightarrow [0,1]$ such that if $x = \sum a_n \frac{1}{3^n}$ for $a_n \in \{0,2\}$, then

$$f(x) = \sum \frac{a_n}{2} \frac{1}{2^n}$$

and we claim that f is a surjection, thus

$$\mathbf{Card}(\mathscr{C}) \geq \mathbf{Card}([0,1])$$

which again shows that $\mathscr{C}$ is uncountable.

Denote $\mathscr{C}$ to be the Cantor Set, if $x \in \mathscr{C}, x = x = \sum_{n=1}^{\infty} a_n \frac{1}{3^n}$ where $a_n \in \{1,2\}$, then we define

$$f(x) = \begin{cases} \sum_{n=1}^{\infty} \frac{a_n}{2} \frac{1}{2^n}, x \in \mathscr{C} \\ \\ f(a), x \notin \mathscr{C}, x \in (a,b) \end{cases}$$

This function is called the *Cantor Lebesgue Function*, here is the graph of $f(x)$:

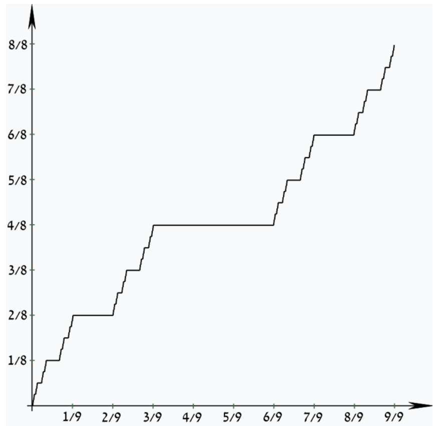


Figure 1.1: Cantor-Lebesgue Function

Facts about $f(x)$:

① $f : [0,1] \longrightarrow [0,1]$

② $f(0) = 0, f(1) = 1, f(\frac{1}{3}) = f(\frac{2}{3}) = \frac{1}{2} \cdots$

③ $f(x)$ is non-decreasing and a surjection.

④ $f(x)$ is continuous.

1.8.2 A subset of $\mathbb{R}$ that is non-measurable : Vitali Set

Then, we have another question: Is there an $A \subseteq \mathbb{R}$, such that A is non-measurable?

The answer is still yes. To see this, we will state the *Axioms of Choice*.

Theorem

Theorem 7. *(Axioms of Choice (AC)):*
If Ω is a collection of non-empty sets, then there exists a function

$$\mathscr{S} : \Omega \longrightarrow \bigcup_{A \in \Omega} A$$

such that

$$\forall A \in \Omega, \mathscr{S}(A) \in A$$

and we call $\mathscr{S}$ to be a selection function. We refer to $\mathscr{S}(A)$ as a representative of A.

Now we will construct a non-measurable set:

Consider $[0,1]$ again, define an equivalence relation $\sim$ on $[0,1]$ given by

$$\forall a,b \in [0,1] : a \sim b \Longleftrightarrow a-b \in \mathbb{Q}$$

Denote by E_a, to be the equivalent class containing a, set Ω to be the collection of all equivalence classes, of course $\forall E_a \in \Omega, E_a \neq \varnothing$.

By *Axioms of Choice*, we can select **EXACTLY ONE** element S_a from E_a for each $E_a \in \Omega$, and set

$$\mathscr{N} := \{S_a : S_a \text{ is a representative of } E_a, E_a \in \Omega\}$$

Proposition

Proposition 30. *We claim that $\mathscr{N}$ is not measurable.*

Proof. Assume $\mathscr{N}$ is measurable instead, then consider $[-1,1]\bigcap\mathbb{Q}$, and $\{q_k : k \geq 1\}$ is an enumeration of $[-1,1]\bigcap\mathbb{Q}$. For each $k \geq 1$, set

$$\mathscr{N}_k = \mathscr{N} + q_k$$

thus, $\mathscr{N}_k \in \mathscr{M}$ and $m(\mathscr{N}) = m(\mathscr{N}_k)$.

Claim 1 : $\mathscr{N}_k$ are disjoint.

Proof. Assume not, the $\exists q_k, q_l \in \mathbb{Q}, (q_k \neq q_l); S_a, S_b \in \mathscr{N}$ such that $S_a + q_k = S_b + q_l$, meaning

$$S_a - S_b = q_l - q_k \in \mathbb{Q}$$

which yields $S_a \sim S_b, q_k = q_l$, a contradiction. ■

Claim 2 : $[0,1] \subseteq \bigcup_{k=1}^{\infty} \mathscr{N}_k$.

Proof. Take any $x \in [0,1]$, then $x \sim S_a$ for some $S_a \in \mathscr{N}$, meaning $x - S_a \in \mathbb{Q}$. Since $x - S_a \in [-1,1]$, then there exists a k, such that $x - S_a = q_k$, so $x \in \mathscr{N}_k$.

On the other hand, $\bigcup_{k=1}^{\infty} \mathscr{N}_k \subseteq [-1,2]$, since $\mathscr{N}_k$ are disjoint, then

$$[0,1] \subseteq \bigcup_{k=1}^{\infty} \mathscr{N}_k \subseteq [-1,2]$$

meaning

$$1 \leq m\left(\bigcup_{k=1}^{\infty} \mathscr{N}_k\right) \leq 3$$

meaning while, we have

$$m\left(\bigcup_{k=1}^{\infty} \mathscr{N}_k\right) = \sum_{k=1}^{\infty} m(\mathscr{N}_k)$$

Since each $\mathscr{N}_k$ has the same measure, it yields a contradiction. ■

That is, $\mathscr{N}$ is not measurable. ■

We call the set

$$\mathscr{N} := \{S_a : S_a \text{ is a representative of } E_a, E_a \in \Omega\}$$

the *Vitali Set*.

Theorem

Theorem 8. *$\forall A \in \mathscr{M}$ with $m(A) > 0$, there exists a subset $B \subseteq A$ such that B is not measurable.*

Proof. Assume otherwise, that is, there exists an $A \in \mathscr{M}$ such that $\forall B \subseteq A$, B is measurable. Let

$$A \subseteq \bigcup_{n \in \mathbb{Z}} (A \cap [n, n+1])$$

so there exists n such that $m(A \bigcap [n, n+1]) > 0$, and $m(A \bigcap [n, n+1] - n) > 0$ because of translation invariant. Let

$$A' := A \cap [n, n+1] - n$$

thus $A' \subseteq [0,1]$ and $m(A') > 0$.

Now consider $\forall B' \subseteq A'$, $B' + n \subseteq A$ and we know that $B' \in \mathscr{M}$. Let $\mathscr{N}, q_k, \mathscr{N}_k$ be the same as earlier defined, let $A'_k = A' \bigcap \mathscr{N}_k$ where A'_k are disjoint and

$$A' = [0,1] \cap A' \subseteq \bigcup_{k=1}^{\infty} (\mathscr{N}_k \cap A') = \bigcup_{k=1}^{\infty} A'_k.$$

Since $m(A') > 0$, so there exists a k such that $m(A'_k) > 0$, by fixing k, set

$$\mathscr{L} := \{l \geq 1 : q_l + q_k \in [-1.1]\}$$

where $\mathscr{L}$ is countably infinite. Consider the set $\{q_l + A'_k, l \in \mathscr{L}\}$:

Since $q_l + q_k \in \mathbb{Q}$, it means that $q_l + q_k = q_m$ for some unique m, then

$$q_l + A'_k = q_l + A' \cap (\mathscr{N} + q_k) \subseteq \mathscr{N}_m$$

also $\{q_l + A'_k, l \in \mathscr{L}\}$ is a family of disjoint sets. Again,

$$\bigcup_{l \in \mathscr{L}} q_l + A'_k \subseteq [-1, 2]$$

$$\sum_{l \in \mathscr{L}} m(q_l + A'_k) \leq 3$$

meaning that $m(q_l + A'_k) = 0$, that is, $m(A'_k) = 0$, a contradiction. ■

1.8.3 A measurable set that is not a Borel set

Still, we ask the following question:

Is there a set $A \in \mathscr{M}$, such that $A \notin \mathfrak{B}_{\mathbb{R}}$?

The answer is again yes.

Let $f : [0,1] \longrightarrow [0,1]$ be the Cantor-Lebesgue function, define

$$g(x) = f(x) + x$$

then $g(x)$ is continuous and strictly increasing, note that

$$g : [0,1] \longrightarrow [0,2]$$

is also a bijection (by I.V.T), then g has an inverse, where

$$g^{-1} : [0,2] \longrightarrow [0,1]$$

exists and also yields a bijection. We say g is a *homeomorphism* (It's not homomorphism!), meaning g maps from open to open; closed to closed sets. So if $A \in \mathfrak{B}_{\mathbb{R}}$, then $g(A) \in \mathfrak{B}_{\mathbb{R}}$.

Observe that if (a,b) is an open interval that gets removed from the construction of the Cantor set, then f is constant on (a,b), then g will map (a,b) to another interval of the length $(b-a)$.

Consider $g([0,1]\backslash\mathscr{C})$, where

$$m(g([0,1]\backslash\mathscr{C})) = 1$$

thus $m(g(\mathscr{C})) = 2 - 1 = 1 > 0$ and $g(\mathscr{C}) \subseteq [0,2]$.

Then there exists $B \subseteq g(\mathscr{C})$ such that $B \notin \mathscr{M}$, let $A = g^{-1}(B)$, then $A \subseteq g^{-1}(g(\mathscr{C})) = \mathscr{C}$.

Since $m(\mathscr{C}) = 0$, then $A \in \mathscr{M}$ and $m(A) = 0$ and $A \notin \mathfrak{B}_{\mathbb{R}}$. Because if so, then $B = g(A) \in \mathfrak{B}_{\mathbb{R}}$ which contradicts the fact that $B \notin \mathscr{M}$.

Chapter 2

Integration Theory

Measurable Functions

We will consider the function f defined on $\mathbb{R}$ that could take $+\infty, -\infty$ as its value, i.e

$$f : \mathbb{R} \longrightarrow \overline{\mathbb{R}}$$

where $\overline{\mathbb{R}} = \mathbb{R} \bigcup \{+\infty, -\infty\}$. We say f is extended real valued.

If $\forall x \in \mathbb{R}, -\infty < f < +\infty$, then we say f is finite valued.

Definition

Definition 12. *Given a function f, take $a \in \mathbb{R}$, we define*

$$f^{-1}([-\infty, a)) := \{x \in \mathbb{R}, f(x) \in [-\infty, a)\}$$

as the inverse image (or pre-image) of f.

Similarly, we have

$$f^{-1}((a, \infty]) := \{x \in \mathbb{R}, f(x) \in (a, +\infty]\}.$$

And generally,

$$\forall B \subseteq \mathbb{R}, f^{-1}(B) := \{x \in \mathbb{R}, f(x) \in B\}$$

Proposition

Proposition 31. ① $\forall B \subseteq \mathbb{R}$, $f^{-1}(B^C) = (f^{-1}(B))^C$.

② $\forall A, B \subseteq \mathbb{R}$, $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$ *and* $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$, *the same also holds for countable union and intersection.*

Theorem

Theorem 9. *Say* $f : \mathbb{R} \longrightarrow \mathbb{R}$ *is measurable, if and only if* $\forall A \subseteq \mathbb{R}$, $f^{-1}([-\infty, a))$ *is measurable.*

We also have these equivalent definitions:

Say $f : \mathbb{R} \longrightarrow \mathbb{R}$ is measurable, if and only if $\forall a \in \mathbb{R}$, we have one of the following: ① $f^{-1}((a, +\infty])$ is measurable; ② $f^{-1}([a, +\infty])$ is measurable; ③ $f^{-1}([-\infty, -a])$ is measurable.

The idea is similar to Borel set, we shall take ③ as an example to illustrate this:

Proof. $\forall a \in \mathbb{R}$, we have

$$f^{-1}([-\infty, a)) = \bigcup_{n=1}^{\infty} f^{-1}\left(\left[-\infty, a - \frac{1}{n}\right]\right)$$

and

$$f^{-1}([-\infty, a]) = \bigcup_{n=1}^{\infty} f^{-1}\left(\left[-\infty, a + \frac{1}{n}\right)\right)$$

■

Also, if f is finite-valued, then f is measurable if and only if $\forall a, b \in \mathbb{R}, a < b$, we have one of the following: ① $f^{-1}((a, b))$ is measurable; ② $f^{-1}([a, b])$ is measurable; ③ $f^{-1}([a, b))$ is measurable; ④ $f^{-1}((a, b])$ is measurable.

Now consider the Borel σ-algebra $\mathfrak{B}_{\overline{\mathbb{R}}}$ of the completion of $\mathbb{R}$ defined by

$$\mathfrak{B}_{\overline{\mathbb{R}}} := \sigma(\mathfrak{B}_{\mathbb{R}} \cup \{+\infty\} \cup \{-\infty\})$$

We claim that $\mathfrak{R}_{\overline{\mathbb{R}}} = \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$.

Proof. $\forall a \in \mathbb{R}$, write $[-\infty, a)$ as $(-\infty, a) \bigcup \{-\infty\}$ and clearly $\sigma((-\infty, a) \bigcup -\infty) \subseteq \mathfrak{B}_{\overline{\mathbb{R}}}$, we only need to prove the other direction.

Note that

$$\{-\infty\} = \bigcap_{n=1}^{\infty} [-\infty, -n), \{+\infty\} = \overline{\mathbb{R}} \Big\backslash \left(\bigcup_{n=1}^{\infty} [-\infty, n)\right)$$

which implies

$$\{+\infty, -\infty\} \in \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$$

thus $\forall a \in \mathbb{R}$, $(-\infty, a) = [-\infty, a) \backslash -\infty \in \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$, meaning

$$\mathfrak{B}_{\overline{\mathbb{R}}} \subseteq \sigma(\{[-\infty, a) : a \in \mathbb{R}\})$$

as desired.

■

Proposition

Proposition 32. *A function $f : \mathbb{R} \longrightarrow \overline{\mathbb{R}}$ is measurable if and only if $\forall B \in \mathfrak{B}_{\overline{\mathbb{R}}}$, $f^{-1}(B) \in \mathscr{M}$.*

Proof. $(\Longleftarrow)$: Obvious, because $[-\infty, a) \in \mathfrak{B}_{\mathbb{R}}, \forall a \in \mathbb{R}$.

$(\Longrightarrow)$: We will use one of the results:

Given a collection of $\mathbb{C}$ of subsets of $\overline{\mathbb{R}}$. we have

$$f^{-1}(\mathscr{C}) := \{f^{-1}(B) : B \in \mathscr{C}\}$$

then $f^{-1}(\sigma(\mathscr{C}) = \sigma(f^{-1}(\mathscr{C}))$.

Now take $\mathscr{C} := \{[-\infty, a) : a \in \mathbb{R}\}$, then

$$f^{-1}(\mathfrak{B}_{\overline{\mathbb{R}}} = f^{-1}(\sigma(\{[-\infty, a) : a \in \mathbb{R}\})) = \sigma(f^{-1}([-\infty, a)) : a \in \mathbb{R}\} \in \mathscr{M}$$

■

Similarly, if f is finite valued, then f is measurable if and only if $\forall B \in \mathfrak{B}_{\mathbb{R}}, f^{-1}(B) \in \mathscr{M}$.

Proposition

Proposition 33. *Given $f : \mathbb{R} \longrightarrow \overline{\mathbb{R}}$, define*

$$f_{\mathbb{R}}(x) = \begin{cases} f(x), if -\infty < f(x) < +\infty \\ \\ 0, otherwise \end{cases},$$

then f is measurable if and only if $\forall B \in \mathfrak{B}_{\mathbb{R}}$, we have $f_{\mathbb{R}}^{-1}(B) \in \mathscr{M}$ and $\{f = +\infty\}, \{f = -\infty\} \in \mathscr{M}$.

Proof. $(\Longleftarrow)$: We assume $R.H.S$, then $\forall a \in \mathbb{R}$, $f^{-1}([-\infty, a)) = \{f = -\infty\} \bigcup f_{\mathbb{R}}^{-1}((-\infty, a)) \in \mathscr{M}$.

$(\Longrightarrow)$: We assume f is measurable, obviously $\{f = +\infty\}, \{f = -\infty\} \in \mathscr{M}$, then $\forall B \in \mathfrak{B}_{\mathbb{R}}$, we have

$$\begin{aligned} f_{\mathbb{R}}^{-1}(B) &= \{x \in \mathbb{R} : f_{\mathbb{R}}(x) \in B\} \\ &= \{x \in \mathbb{R}, f_{\mathbb{R}}(x) \in B, -\infty < f(x) < +\infty\} \cup \{x \in \mathbb{R}, f(x) = \pm\infty\} \in \mathscr{M} \end{aligned}$$

■

Definition

Definition 13. *If a statement is true for every $x \in A$ for $A \in \mathscr{M}$ such that $m(A^C) = 0$, then we say that the statement is true almost everywhere (is true for almost every x).*

Proposition

Proposition 34. *If $f : \mathbb{R} \longrightarrow \overline{\mathbb{R}}$ is measurable and $f = g$ almost everywhere, then g is measurable.*

Corollary

Corollary 3. *If f is a finite valued almost everywhere, then f is measurable if and only if $f_{\mathbb{R}}$ is measurable, if and only if $\forall a, b \in \mathbb{R}, a < b$ we have $f^{-1}((a,,b)) \in \mathscr{M}$.*

Proposition

Proposition 35. ① *If $f(x) = c$ (f is constant function), then f is measurable;*

② *If $f = \mathbf{1}_A$ for some $A \subseteq \mathbb{R}$, then f is measurable if and only if $A \in \mathscr{M}$, where f is the characteristic function of A defined by*

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

Proof. ① : Assume $f(x) = c$, then we have $\forall a \in \mathbb{R}$,

$$f^{-1}([-\infty, a)) = \begin{cases} \mathbb{R}, c < a \\ \varnothing, c \geq a \end{cases} \quad \in \mathscr{M}$$

② : Assume $f = \mathbf{1}_A$, then we have $\forall a \in \mathbb{R}$,

$$f^{-1}([-\infty, a)) = \begin{cases} \mathbb{R}, a > 1 \\ \varnothing, a \leq 0 \\ A^C, 0 < a \leq 1 \end{cases} \quad \in \mathscr{M}$$

if and only if $A \in \mathscr{M}$. ■

Proposition

Proposition 36. *If f is a finite valued continuous function on $\mathbb{R}$, then f is measurable.*

Proof. The fact that if $f : \mathbb{R} \longrightarrow \mathbb{R}$ is continuous, if and only if for all open set $G \subseteq \mathbb{R}$, $f^{-1}(G)$ is open. Then $\forall a, b \in \mathbb{R}$, $f^{-1}((a, b))$ is open $\in \mathscr{M}$. ■

In fact, if $f : \mathbb{R} \longrightarrow \mathbb{R}$ is continuous, then $\forall B \in \mathfrak{B}_{\mathbb{R}}$, $f^{-1}(B) \in \mathfrak{B}_{\mathbb{R}}$.

Further, if f^{-1} exists and is continuous, then $\forall B \in \mathfrak{B}_{\mathbb{R}}, f(B) \in \mathfrak{B}_{\mathbb{R}}$.

Proposition

Proposition 37. *If $f : \mathbb{R} \longrightarrow \mathbb{R}$ (finite-valued) is measurable and $g : \mathbb{R} \longrightarrow \mathbb{R}$ is continuous, then $g \circ f : \mathbb{R} \longrightarrow \mathbb{R}$ is measurable. (Note that the order matters!)*

Proof. $\forall a \in \mathbb{R}$, we have

$$(g \circ f)^{-1}([-\infty, a)) := \{x \in \mathbb{R}, g(f(x)) < a\}$$

$$= \{x \in \mathbb{R}, f(x) \in g^{-1}([-\infty, a))\}$$

$$= f^{-1}(g^{-1}([-\infty, a))) \in \mathscr{M}$$

∎

Proposition

Proposition 38. *If a function $f : \mathbb{R} \longrightarrow \overline{\mathbb{R}}$ is measurable, then*

① $\forall c \in \mathbb{R}$, cf *is measurable;*

② $|f|$ *is measurable;*

③ $\forall k \in \mathbb{N}$, f^k *is measurable.*

④ *If f, g are two measurable functions (finite-valued), then $f + g$, $f \cdot g$, $f \vee g := \max\{f, g\}$, $f \wedge g := \min\{f, g\}$ are all measurable functions.*

We will only give a proof on ④:

Proof. $\forall a \in \mathbb{R}$, we have

$$(f+g)^{-1}([-\infty, a)) = \{x \in \mathbb{R}, f(x) + g(x) < a\}$$

$$= \{x \in \mathbb{R}, f(x) < a - g(x)\}$$

$$= \bigcup_{q \in \mathbb{Q}} \{x \in \mathbb{R}, f(x) < q < a - g(x)\}$$

$$= \bigcup_{q \in \mathbb{Q}} (\{x \in \mathbb{R}, f(x) < q\} \cap \{x \in \mathbb{R}, g(x) < a - q\}) \in \mathscr{M}$$

Thus $f-g$ is also measurable, and we have

$$f\cdot g=\frac{1}{4}[(f+g)^2-(f-g)^2]$$

is also measurable, and

$$f\vee g=\frac{1}{2}(|f-g|+|f+g|$$

is also measurable, and finally

$$f\wedge g=-\max\{-f,-g\}$$

is also measurable. ∎

Corollary

Corollary 4. *If a function f is measurable, then $f^+ := f\vee 0=\max\{f,0\}$ and $f^-=-(f\vee 0)=\max\{-f,0\}$ are also measurable.*

Proposition

Proposition 39. *Let $\{f_n : n\geq 1\}$ be a sequence of measurable functions, then*

$$\sup(f_n),\inf(f_n),\lim_{n\to\infty}\sup(f_n),\lim_{n\to\infty}\inf(f_n)$$

are all measurable functions.

Proof. To show that $\sup(f_n)$ is measurable, we will study that $\forall a\in\mathbb{R},\{\sup f_n\leq a\}$.

$x\in\{\sup f_n\leq a\}$ implies $\sup f_n(x)\leq a$, meaning $f_n(x)\leq a,\forall n$. Thus $x\in\bigcup_{n=1}^{\infty}\{f_n\leq a\}$, that is

$$\{\sup f_n\leq a\}=\cup_{n=1}^{\infty}\{f_n<a\}\in\mathscr{M},$$

meaning $\sup f_n$ is measurable. ∎

Proposition

Proposition 40. *Let $\{f_n : n\geq 1\}$ be a sequence of measurable functions. then all of the following sets are measurable:*

① $\{x\in\mathbb{R} : \lim f_n(x)$ *exists*$\}$;

② $\{\lim f_n=\pm\infty\};\{\lim f_n=c : c\in\mathbb{R}\}$.

Proof of ① :

Proof. ① $\iff\{\limsup f_n=\liminf f_n\}$, and finite. Thus ① $=\{-\infty<\limsup f_n<+\infty\}\cap\{-\infty<\liminf f_n<+\infty\}\cap\{\limsup f_n=\liminf f_n\}$, where everything is measurable. ∎

Proof of ② :

Proof. First consider $\{\lim f_n = c\}$, meaning $\forall x \in \mathbb{R}$, the set

$$\{\forall \varepsilon > 0, \exists n \geq 1 : \forall m \geq n, |f_m(x) - c| \leq \varepsilon\}$$

can be written as

$$\{\forall x \in \mathbb{R}, k \geq 1, \exists n \geq 1 : \forall m \geq n, |f_m(x) - c| \leq \frac{1}{k}\}.$$

that is,

$$\bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{|f_m(x) - c| \leq \frac{1}{k}\} \in \mathscr{M}.$$

For the other part, use the definition of a *Cauchy Sequence*. ■

Approximation by simple functions

Given a function $f : \mathbb{R} \longrightarrow \overline{\mathbb{R}}$ is measurable, then:

① $f = f^+ - f^-$, where $f^+, f^- \geq 0$ and are all measurable.

② $\forall n \geq 1$, $f_n^+ = (f^+ \wedge n)\mathbf{1}_{[-n,n]}$.

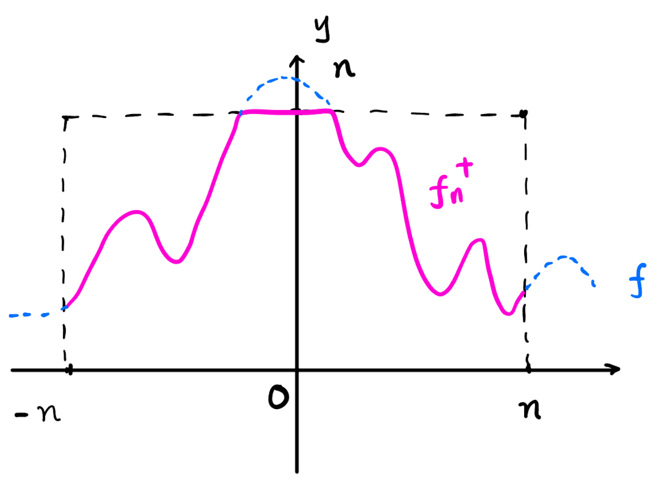


Figure 2.1: The graph of f_n^+, where $f^+ \wedge n)$ represents a truncation, and $\mathbf{1}_{[-n,n]}$ represents a cut-off

Obviously, f_n^+ is non-negative and measurable, bounded by n, zero outside of a bounded set, but also $\{f_n^+ : n \geq 1\}$ is increasing and $\lim f_n^+ = f^+$.

Similarly, define $f_n^- = (f^- \wedge n)\mathbf{1}_{[-n,n]}$, still $\{f_n^- : n \geq 1\}$ is increasing and $\lim f_n^- = f^-$.

③ Fix n, consider f_n^+ and for $k = 0, 1, 2, \cdots, n2^n$, define

$$A_{n,k} := \left\{x \in [-n,n] : \frac{k}{2^n} \leq f_n^+(x) < \frac{k+1}{2^n}\right\} \in \mathscr{M}$$

We claim that $A_{n,k} \bigcap A_{n,l} = \varnothing$, if $k \neq l$, then set

$$\varphi_n = \sum_{k=0}^{n2^n} \mathbf{1}_{A_{n,k}} \frac{k}{2^n}$$

where these φ_n's are simple functions.

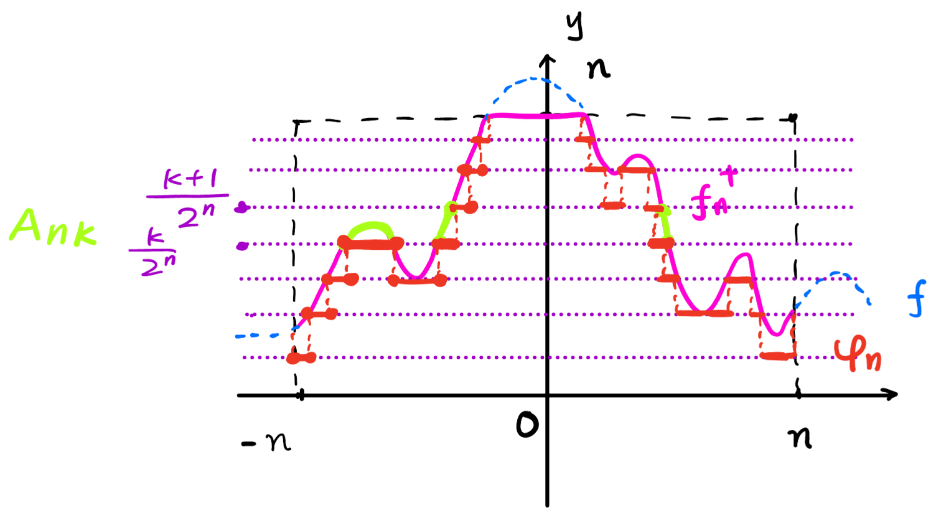


Figure 2.2: A more detailed picture on the functions we defined.

Definition

Definition 14. *φ is a simple function if*

$$\varphi = \sum_{k=1}^{L} \mathbf{1}_{E_k} a_k$$

where L is a positive integer, a_k's are constant and $E_k \in \mathscr{M}$, $m(E_K) < +\infty$.

Also, we claim that $\{\varphi_n : n \geq 1\}$ is increasing, and $\lim \varphi_n = f^+$.

Proof. We will prove that $\lim_{n\to\infty} \varphi_n(x) = f^+(x)$.

$\forall n \in \mathbb{R}$, when n is large enough, $x \in [-n,n]$ will hold. Then clearly $f^+(x) = f^+(x)\mathbf{1}_{[-n,n]}(x)$. Assume $f^+(x) < +\infty$, then by making n even larger when necessary, one can make $f^+(x) < n$. In fact, $f^+(x) < n \implies f^+(x) = (f \wedge n)\mathbf{1}_{[-n,n]}(x) = f_n^+(x)$. Further, we observe that $0 \leq f_n^+(x) - \varphi_n(x) < 2^{-n}$, meaning that $0 \leq f^+(x) - \varphi_n(x) < 2^{-n}$, as $n \to \infty$, we have $\lim_{n\to\infty} \varphi_n(x) = f^+(x)$.

Now assume that $f^+(x) = +\infty$, then it must be that $\varphi_n(x) = n$ for all large n, thus it is still true. ■

Theorem

Theorem 10. *If g is measurable and $g \geq 0$, then there exists a sequence of simple functions $\{\varphi_n : n \geq 1\}$ such that $\{\varphi_n : n \geq 1\}$ is increasing and $\lim_{n\to\infty} \varphi_n(x) = g(x)$ for all $x \in \mathbb{R}$.*

Theorem

Theorem 11. *If f is measurable, then there exists a sequence of simple functions $\{\xi_n : n \geq 1\}$ such that $|\xi_n|$ is increasing and $|\xi_n| \leq |f|$ for all $x \in \mathbb{R}$ and $\lim_{n\to\infty} \xi_n(x) = f(x)$.*

Now, we would like to approximate the function by step functions.

Definition

Definition 15. *θ is called a step function, if*

$$\theta(x) = \sum_{k=1}^{L} a_k \mathbf{1}_{I_k}(x)$$

where $L \in \mathbb{N}$ and a_k's are constant, I_k's are finite open intervals.

Theorem

Theorem 12. *If f is measurable, then there exists a sequence of step functions $\{\theta_n : n \geq 1\}$ such that*

$$\lim_{n\to\infty} \theta_n(x) = f(x)$$

for almost every $x \in \mathbb{R}$.

Proof. *W.L.O.G*, assume f is non-negative. Given that $A \in \mathscr{M}$ with finite measure, recall that $\forall \varepsilon > 0$, $\exists I_1, I_2, \cdots, I_n$ open and finite, such that $m(A \triangle \bigcup I_n) < \varepsilon$. By re-arranging these I_i's if necessary, we may assume that I_i's are pairwise disjoint, hence

$$\mathbf{1}_{\bigcup I_i} = \sum_{i=1}^{n} \mathbf{1}_{I_i}$$

Now define

$$\theta_A = \sum_{n=1}^{n} \mathbf{1}_{I_i},$$

then we know that

$$m(\{x \in \mathbb{R} : \mathbf{1}_A(x) \neq \theta_A\}) < \varepsilon$$

where

$$\{x \in \mathbb{R} : \mathbf{1}_A(x) \neq \theta_A\} = A \triangle \left(\bigcup I_n\right)$$

Since f is measurable and non-negative, then there exists a sequence of simple functions $\{\varphi_n : n \geq 1\}$, such that $\lim_{n\to\infty} \varphi_n = f$, in particular,

$$\varphi_n = \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbf{1}_{A_{n,k}}$$

It means that $\forall n \geq 1; k = 0, 1, 2, \cdots, n2^n$, we can find a step function $\theta_{n,k}$ such that

$$m\left(\{x \in \mathbb{R} : \mathbf{1}_{A_{n,k}}(x) \neq \theta_{n,k}(x)\}\right) < \frac{1}{2^n(n2^n + 1)}$$

Set

$$\theta_n := \sum_{k=0}^{n2^n} \frac{k}{2^n}\theta_{n,k}; E_n := \{x \in \mathbb{R}, \theta_n(x) \neq \varphi_n(x)\}$$

Then

$$\begin{aligned} m(E_n) &\leq m\left(\bigcup_{k=0}^{n2^n}\{\theta_{n,k} \neq \mathbf{1}_{A_{n,k}}\}\right) \\ &\leq \sum_{k=0}^{n2^n} m\left(\{\theta_{n,k} \neq \mathbf{1}_{A_{n,k}}\}\right) \\ &\leq (n2^n+1)\frac{1}{2^n(n2^n+1)} \\ &= 2^{-n} \end{aligned}$$

It implies that $m(E_n) \leq 2^{-n}$, φ_n's are chosen such that $\forall x \in \mathbb{R}, |\varphi_n(x) = f_n(x)| \leq 2^{-n}$, if $F_n := \{x \in \mathbb{R} : |\theta_n(x) - f_n(x)| > 2^{-n}\}$, then $F_n \subseteq E_n$ and $m(F_n) \leq 2^{-n}$.

Corollary

Corollary 5. *For almost every $x \in \mathbb{R}$, there exists $m \geq 1$, such that $\forall n \geq m$,*

$$|\theta_n(x) - f_n(x)| \leq \frac{1}{2^n}$$

Proof. Consider the compliment:

$\forall m$, such that $\exists n \geq m$, we have $|\theta_n(x) - f_n(x)| > 2^{-n}$. In terms of sets operation, we will show that

$$m\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{x \in \mathbb{R} : |\theta_n(x) = f_n(x)| > \frac{1}{2^n}\}\right) = 0$$

i.e $m(\bigcap\bigcup F_n) = 0$.

Define $B_m = \bigcup_{n=m}^{\infty} F_n$, clearly B_m is decreasing, then

$$m\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} F_n\right) = \lim_{m\to\infty} m(B_m) \leq \lim_{m\to\infty}\sum_{n=m}^{\infty} m(F_n) = 0$$

■

Then for almost everywhere, $\exists m \geq 1$, $\forall n \geq m, |\theta_n - f_n| \leq 2^{-n}$. Meaning that fora almost everywhere, $\lim_{n\to\infty}(\theta_n - f_n) = 0$, therefore, we have $\theta_n = f$ as $n \to \infty$ almost everywhere.

■

Our previous work will result in an important lemma:

Theorem

Theorem 13. *(Borel-Cantelli)*

If there exists a sequence of measurable intervals $\{F_n : n \geq 1\} \subseteq \mathscr{M}$ *such that* $\sum_{n=1}^{\infty} m(F_n) < +\infty$, *then* $m(\limsup_n F_n) = 0$, *i.e*

$$m\left(\bigcap_{n=1}^{\infty} \bigcup_{n=m}^{\infty} F_n\right) = 0$$

Convergence Almost Everywhere and Convergence in Measure

We begin by introducing the definitions:

Definition

Definition 16. *Suppose* $\{f_n : n \geq 1\}$ *and* f *are all measurable functions, we say the sequence* $\{f_n : n \geq 1\}$ *convergence to* f *almost everywhere and write* $f_n \to f$ *if for almost every* $x \in \mathbb{R}$, $\lim_{n\to\infty} f_n(x) = f(x)$.

Definition

Definition 17. *Suppose* $\{f_n : n \geq 1\}$ *and* f *are all measurable functions, we say the sequence* $\{f_n : n \geq 1\}$ *convergence to* f *in measure and write* $f \to f$ *in measure if*

$$\forall \delta > 0, \lim_{n\to\infty} m(\{x \in \mathbb{R} : |f_n(x) - f(x)| \geq \delta\}) = 0$$

The relation between convergence in measure and convergence almost everywhere can be described by the following proposition:

Proposition

Proposition 41. *Given a sequence of finite-valued measurable functions* $\{f_n : n \geq 1\}$ *and* f, *and* $A \in \mathscr{M}$ *where* $m(A) < +\infty$. *If* $f_n \to f$ *almost everywhere on* A, *then* $f_n \to f$ *in measure on* A.

The converse is not true!

Proof. $\forall\delta > 0$,

$$
\begin{aligned}
&\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{x\in A : |f_n(x)-f(x)|>\delta\} \subseteq \{x\in A, \lim_{n\to\infty} f_n(x)\neq f(x)\}\\
&\implies m\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{x\in A : |f_n(x)-f(x)|>\delta\}\right)=0\\
&\implies \lim_{m\to\infty} m(B_m)=0\\
&\implies \{|f_n(x)-f(x)|>\delta\} \subseteq \bigcup_{n=m}^{\infty}\{|f_n-f|>\delta\}\\
&\implies m(\{|f_m-f|>\delta\}) \leq m(B_m)=0, n\to\infty
\end{aligned}
$$

■

Note that in the proposition, the assumption $m(A) < +\infty$ is necessary.

Proof. Let $f_n = \mathbf{1}_{[n,+\infty)}, f = 0$. Then $\lim_{n\to\infty} f_n(x) = f(x) = 0$, however $m(\{x \in \mathbb{R} : |f_n(x) = f(x)| = 1\} = m([n,+\infty)) = +\infty$, a contradiction. ■

Theorem

Theorem 14. *Given $\{f_n : n \geq 1\}$ and f are measurable, finite-valued functions, if $f_n \to f$ in measure, then there exists a subsequence $\{n_k : k \geq 1\} \subseteq \mathbb{N}$ such that $f_{n_k} \to f$ almost everywhere.*

Proof. Assume $f_n \to f$ in measure, then $\forall\delta > 0$, $m(\{|f_n - f| > \delta\}) \to 0$ as $n \to \infty$, meaning that $\forall k \geq 1$, $\exists n_k$ such that $m(\{f_{n_k} - f| > 1/k\}) \leq 1/k^2$, denote $A_k := \{f_{n_k} - f| > 1/k\}$, then by *Borel-Cantelli Lemma*, we have

$$
m\left(\bigcap_{l=1}^{\infty}\bigcup_{k=l}^{\infty}A_k\right) = \lim_{l\to\infty} m\left(\bigcup_{k=l}^{\infty}A_k\right) \leq \lim_{l\to\infty}\sum_{k=l}^{\infty} m(A_k) = 0,
$$

meaning that for almost everywhere, $\exists l$, such that $\forall k \geq l : |f_{n_k} - f| \leq 1/k$, which means that for almost everywhere, $\lim_{k\to\infty}|f_{n_k} - f| = 0$, thus $f_{n_k} \to f$ almost everywhere. ■

Theorem

Theorem 15. *(Subsequence Test)*
Given $\{f_n : n \geq 1\}$ and f are measurable and finite-valued functions, then $f_n \to f$ in measure if and only if for any subsequence $\{n_k\}$, there exists a sub-subsequence $\{n_{k_l}\} \subseteq \{n_k\} \subseteq \mathbb{N}$ such that $f_{n_{k_l}} \to f$ in measure as $l \to \infty$

Proof. ($\Longrightarrow$) is obvious, now consider ($\Longleftarrow$). We will prove by contradiction. Assume otherwise, that is, assume $f_n \nrightarrow f$ in measure, then $\exists\delta > 0, \{n_k\}$, such that

$$
m(\{|f_{n_k} - f| > \delta\}) > \delta, \forall k
$$

Then by the hypothesis on the right hand side, $\exists\{n_{k_l}\}$ such that $f_{n_{k_l}} \to f$ in measure, which yields a contradiction. ■

Here is an exercise for the reader: Assume all functions are measurable and finite-valued, if $f_n \to f$, $g_n \to g$ in measure, then $f_n g_n \to fg$ in measure.

Egorov's Theorem and Lusin's Theorem

Recall that if f is measurable, then there exists a sequence of step functions $\{\theta_n\}$ such that $\theta_n \to f$ almost everywhere.

Theorem

Theorem 16. *(Egorov's)*
Given $\{f_n\}, f$ are measurable and finite-valued functions, and $A \in \mathscr{M}$ with $m(A) < +\infty$. If $f_n \to f$ almost everywhere on A, then $\forall \varepsilon > 0$, there exists a closed set $A_\varepsilon \subseteq A$ with $m(A \backslash A_\varepsilon) < \varepsilon$ such that $f_n \to f$ uniformly on A_ε.

Proof. We first assume that f is finite-valued on A, we want to show that $\forall \varepsilon > 0$, there exists a closed set A_ε with $m(A \backslash A_\varepsilon) \leq \varepsilon$, we have

$$\sup_{x \in A_\varepsilon} |f_n(x) - f(x)| \to 0, n \to \infty$$

For each $k \geq 1$, (Now we fix k), and each $n \geq 1$, set

$$E_n^{(k)} := \left\{ x \in A : |f_j(x) - f(x)| < \frac{1}{k}, \forall j \geq n \right\},$$

we see that $E_n^{(k)} \subseteq E_{n+1}^{(k)}$, thus

$$\bigcup_{n=1}^{\infty} E_n^{(k)} := \left\{ x \in A : \exists n \geq 1, \forall j \geq n, |f_j(x) - f(x)| \leq \frac{1}{k} \right\}$$

where

$$\{x \in A : \lim_{n \to \infty} f_n(x) = f(x)\} \subseteq \bigcup_{n=1}^{\infty} E_n^{(k)}$$

If we let $A' = \{x \in A : \lim_{n \to \infty} f_n(x) = f(x)\}$, then by hypothesis, we have $m(A') = m(A)$, and thus

$$m(A) = m(A') \leq m\left(\bigcup_{n=1}^{\infty} E_n^{(k)} \right) = \lim_{n \to \infty} m(E_n^{(k)}) \leq m(A)$$

meaning $\lim_{n \to \infty} m(E_n^{(k)}) = m(A)$.

Now given $\forall \varepsilon > 0$, $\exists n_k \geq 1$, such that $m(A \backslash E_{n_k}^{(k)}) = m(A) - m(E_{n_k}^{(k)}) < \frac{1}{2^n} \frac{\varepsilon}{2}$. Also set

$$B := A \Big\backslash \bigcap_{k=1}^{\infty} E_{n_k}^{(k)}$$

then

$$m(B) = m\left(\bigcup_{k=1}^{\infty} A\backslash E_{n_k}^{(k)}\right) \leq \sum_{k=1}^{\infty} m(A\backslash E_{n_k}^{(k)}) \leq \frac{\varepsilon}{2}$$

Now set $G := A\backslash B = \bigcap_{k=1}^{\infty} E_{n_k}^{(k)}$, then if $x \in A$, of course $x \in E_{n_k}^{(k)}$, $\forall k \geq 1$.

Thus $\forall k \geq 1$, $\forall j = n_k, |f_j(x) - f(x)| \leq \dfrac{1}{k}$, meaning that $f_n \to f$ uniformly on G. By the properties of the regularity, there exists a closed A_ε, $A_\varepsilon \subseteq G$ with $m(G\backslash A_\varepsilon) < \varepsilon/2$, thus $f_n \to f$ uniformly on A_ε and $m(A\backslash A_\varepsilon) = m(A\backslash G) + m(G\backslash A_\varepsilon)$, meaning that

$$m(A\backslash A_\varepsilon) = m(B) + m(G\backslash A_\varepsilon) \leq \varepsilon$$

If $f = \pm\infty$, split f into three parts and do the exact same as above. ■

Here are some remarks:

① The assumption that $m(A) < +\infty$ is necessary. For example. we set $f_n = \mathbf{1}_{[n,+\infty)}$, then $f_n \to 0$ pointwisely, but $\forall a \in \mathbb{R}$, f_n does not converge to 0 uniformly.

② In general, Egorov's theorem does not apply that $f_n \to f$ almost everywhere. For example, on $[0,1]$ consider $f_n(x) = x^n$, $f(x) = 0$. $\forall x \in [0,1)$, $f_n(x) \to f(x)$ as $n \to \infty$, and $f_n(x) \to f(x)$ almost everywhere on $[0,1]$. If $A \subseteq [0,1]$ is closed and $1 \in A$, then $f_n \nrightarrow f$ uniformly on A. To see this, take $\{x_m\} \subseteq A$ with $\{x_m\}$ increasing and $\lim x_m = 1$. Consider

$$\sup_{x\in A} |f_n(x) - f(x)| \geq \sup_m |f_n(x_m) - f(x_m)| = \sup_m x_m^n = 1$$

thus $f \nrightarrow f$ uniformly on A. But in fact, $\forall \varepsilon > 0$, $f_n \to f$ uniformly on $[0, 1-\varepsilon]$.

Theorem

Theorem 17. *(Lusin's)*
Given f is measurable, $A \in \mathscr{M}$ with $m(A) < +\infty$, then $\forall \varepsilon > 0$, there exists a closed $A_\varepsilon \subseteq A$ with $m(A\backslash A_\varepsilon) < \varepsilon$ such that

$$f\Big|_{A_\varepsilon} \text{ is continuous}$$

Here is an important remark: Lusin's Theorem says that $f\Big|_{A_\varepsilon}$ as a function on A_ε is continuous, NOT the same as f as a function on A is continuous at points in A_ε. Given $f = \mathbf{1}_{\mathbb{Q}\cap[0,1]}$, we know that f is not continuous everywhere, however $f\Big|_{\mathbb{Q}\cap[0,1]}$ is a constant and therefore continuous on $[0,1]\cap\mathbb{Q}$.

Proof. Recall that there exists a sequence $\{\theta_n\}$ of step functions such that $\theta_n \to f$ almost everywhere on A,, and θ_n is piecewise constant and piecewise continuous.Given $\varepsilon > 0$, it is possible to find E_n open, such that

$$m(E_n) \leq \frac{\varepsilon}{2}\frac{1}{2^n}$$

and

$$\theta_n\Big|_{E_n^C}$$

is continuous. Meanwhile, by Egorov's theorem, we know that $\exists B \subseteq A$ where B is closed, such that $m(A\backslash B) \leq \varepsilon/2$ and $\theta_n \to f$ uniformly on B. We now set

$$A_\varepsilon = B \Bigg\backslash \bigcup_{n=1}^{\infty} E_n$$

and we know that $A_\varepsilon \subseteq A$, and

$$\begin{aligned} m(A\backslash A_\varepsilon) &\leq m(A\backslash B) + m(B\backslash A_\varepsilon) \\ &\leq m(A\backslash B) + m\left(\bigcup_{n=1}^{\infty} E_n\right) \\ &\leq \frac{\varepsilon}{2} + \sum_{n=1}^{\varepsilon} + \sum_{n=1}^{\infty} m(E_n) \leq \varepsilon \end{aligned}$$

Finally, on A_ε, $\theta_n \to f$ uniformly and $\theta_n\Big|_{A_\varepsilon}$ is continuous, implies $f\Big|_{A_\varepsilon}$ is continuous. ■

Here are more remarks:

① Lusin's Theorem does not apply that f is continuous almost everywhere.

② θ_n used in the proof is not continuous on $\mathbb{R}$, but in fact, we can choose a sequence $\{\widetilde{\theta_n}\}$ that is continuous on $\mathbb{R}$ such that $\widetilde{\theta_n} \to f$ almost everywhere.

③ Lusin's theorem implies that $\forall k$ large enough, $\exists$ closed $A_k \subseteq A$ such that $m(A\backslash A_k) \leq 1/k$ and $f\Big|_{A_k}$ is continuous. In fact, we can take A_k to be increasing, otherwise use $\widetilde{A_k} := \bigcup_{i=1}^{k} A_i$ to replace A_k.

Construction of Integrals

2.5.1 Integral of Simple Functions

Definition

Definition 18. *Given a simple function $\varphi = \sum_{k=1}^{L} a_k \mathbf{1}_{E_k}$, then the Lebesgue integral of φ is defined as*

$$\int_{\mathbb{R}} \varphi(x)dx = \sum_{k=1}^{L} a_k m(E_k)$$

Corollary

Corollary 6.

$$\int_{\mathbb{R}} \mathbf{1}_A = m(A), \quad , \forall A \in \mathscr{M}, m(A) < +\infty$$

Now, we have the properties of the Legesgue integral of simple functions:

Proposition

Proposition 42. *(Well Defined)*
If $\varphi = \sum_{k=1}^{L} a_k \mathbf{1}_{E_k} = \sum_{l=1}^{M} b_l \mathbf{1}_{F_l}$, for $E_k, F_l \in M, a_k, b_l \in \mathbb{R}$ then

$$\sum_{k=1}^{L} a_k \mathbf{1}_{E_k} = \sum_{l=1}^{M} b_l \mathbf{1}_{F_l}$$

Proof. *W.L.O.G*, assume E_k's are pairwise disjoint, so are F_l's. Set $a_0 = b_0 = 0$, and define

$$E_0 = \left(\bigcup_{k=1}^{L} E_k\right)^C ; F_0 = \left(\bigcup_{l=1}^{M} F_l\right)^C$$

i.e, $\{E_0, E_1, \cdots, E_L\}; \{F_0, F_1, \cdots, F_M\}$ are both a partition of $\mathbb{R}$. Then $\varphi = \sum_{k=1}^{L} a_k \mathbf{1}_{E_k} = \sum_{k=0}^{L} a_k \mathbf{1}_{E_k}$, and note that $\mathbf{1}_{E_k} = \sum_{l=0}^{M} \mathbf{1}_{E_k \cap F_l}$, meaning

$$\varphi = \sum_{k=0}^{L} \sum_{l=0}^{M} a_k \mathbf{1}_{E_k \cap F_l}$$

Similarly, we have

$$\varphi = \sum_{l=1}^{M} b_l \mathbf{1}_{F_l} = \sum_{l=0}^{M} b_l \mathbf{1}_{F_l} = \sum_{l=0}^{M} \sum_{k=0}^{L} b_l \mathbf{1}_{F_l \cap E_k}$$

If $E_k \cap F_l \neq \varnothing$, then $a_l = b_l$, meaning that

$$\int_{\mathbb{R}} \varphi = \sum_{k=1}^{L} a_k m(E_k) = \varphi = \sum_{k=0}^{L} \sum_{l=0}^{M} a_k m(E_k \cap F_l)$$

$$\int_{\mathbb{R}} \varphi = \sum_{k=1}^{L} a_k m(E_k) = \varphi = \sum_{l=0}^{M} \sum_{k=0}^{L} b_l m(E_k \cap F_l)$$

If $E_k \cap F_l = \varnothing$, then $m(E_k \cap F_l) = 0$, which is enough to match. ■

Proposition

Proposition 43. *(Linearity)*
If φ and ω are two simple functions, $a, b \in \mathbb{R}$, then $a\varphi + b\omega$ is again a simple function and

$$\int_{\mathbb{R}} a\varphi + b\omega = a \int_{\mathbb{R}} \varphi + b \int_{\mathbb{R}} \omega$$

Proof. Obvious ■

Proposition

Proposition 44. *(Finite Additivity)*
If φ is a simple function, and $A_1, A_2, \cdots, A_N \in \mathscr{M}$ where $m(A_i \cap A_j) = \varnothing, i \neq j$, then

$$\int_{\cup_{n=1}^{N} A_n} \varphi = \sum_{n=1}^{N} \int_{A_n} \varphi$$

Proof. Obvious ■

Note that $\forall A \in \mathscr{M}$, $\mathbf{1}_A \varphi$ is again a simple function, and

$$\int_A \varphi = \int_{\mathbb{R}} \mathbf{1}_A \varphi$$

Proposition

Proposition 45. *(Monotoncity)*
If φ, ω are simple functions with $\varphi \leq \omega$, then

$$\int_{\mathbb{R}} \varphi \leq \int_{\mathbb{R}} \omega$$

Proof. Assume $\varphi = \sum_{k=1}^{L} a_k \mathbf{1}_{E_k}$, $\omega = \sum_{l=1}^{M} b_l \mathbf{1}_{F_l}$, then from proposition 42, we may easily see that it's true. ■

Proposition

Proposition 46. *If φ is a simple function, then $|\varphi|$ is also a simple function, and*

$$\left| \int_{\mathbb{R}} \varphi \right| \leq \int_{\mathbb{R}} |\varphi|$$

2.5.2 Integral of non-negative functions

Definition

Definition 19. *If f is a non-negative, measurable functions, then the Lebesgue integral of f is defined as*

$$\int_{\mathbb{R}} f(x)dx := \sup \left\{ \int_{\mathbb{R}} h(x)dx : h \text{ is simple}, h \leq f \right\}$$

Theorem

Theorem 18. *This definition is the same as definition 16 restricted to non-negative simple functions.*

Proof. Assume $\varphi = \sum_{k=1}^{L} a_k \mathbf{1}_{E_k}$ is non-negative, thus for f non-negative, we should have

$$\int_{\mathbb{R}} f \geq \int_{\mathbb{R}} \varphi$$

which means we only need to show that

$$\int_{\mathbb{R}} f \leq \int_{\mathbb{R}} \varphi$$

We will show that $\forall \omega$ simple functions, $\omega \leq \varphi$, meaning $\int_{\mathbb{R}} \omega \leq \int_{\mathbb{R}} \varphi$ by monotoncity. ■

Notice that given $f \geq 0$ to be measurable, there exists a sequence of simple functions $\{\varphi_n : n \geq 1\}$ such that $\varphi_n \leq f$ and

$$\lim_{n\to\infty} \int_{\mathbb{R}} \varphi_n = \int_{\mathbb{R}} f$$

Theorem

Theorem 19. *Suppose $f \geq 0$ and is measurable, if $\{\varphi_n\}$ is a sequence of simple functions such that $\varphi_n \leq \varphi_{n+1}$, and $\lim_{n\to\infty} \varphi_n = f$ pointwise, then*

$$\lim_{n\to\infty} \int_{\mathbb{R}} \varphi_n = \int_{\mathbb{R}} f$$

Proof. $L.H.S = \lim\limits_{n\to\infty} \int_{\mathbb{R}\varphi_n}$, $R.H.S = \int_{\mathbb{R}} f$, since $\lim_{n\to\infty} \varphi_n = f$, thus by definition $L.H.S \leq R.H.S$ holds trivially. So we only need to show

$$\lim_{n\to\infty} \int_{\mathbb{R}} \varphi_n \geq \int_{\mathbb{R}} f$$

i.e by showing $\forall \omega$ simple, $\omega \leq f$, we have

$$\int_{\mathbb{R}} \omega \leq \lim_{n\to\infty} \int_{\mathbb{R}} \varphi_n$$

Assume $\omega = \sum_{k=1}^{L} a_k \mathbf{1}_{E_k} = \sum_{k=0}^{L} a_k \mathbf{1}_{E_k}$, where $\{E_0, E_1, \cdots, E_L\}$ is a partition of $\mathbb{R}$. Since

$$\int_{\mathbb{R}} \omega = \sum_{k=0}^{L} a_k m(E_k), \int_{\mathbb{R}} \varphi = \sum_{k=0}^{L} \int_{E_k} \varphi_n,$$

it's sufficient to show that $\forall k = 0, 1, \cdots, L$,

$$a_k m(E_k) \leq \lim_{n\to\infty} \int_{E_k} \varphi_n \qquad (*)$$

Fix k for now, if $a_k = 0$ or $m(E_k) = 0$, then $(*)$ holds trivially.

Assume $a_k > 0, m(E_k) > 0$, then for a fixed k,

$$\lim_{n\to\infty} \varphi_n = f \geq \omega \implies \forall x \in E_k, \lim_{n\to\infty} \varphi_n(x) \geq a_k$$

Now $\forall \varepsilon > 0$, set

$$C_n^{\varepsilon} := \{x \in E_K : \Phi_n(x) \geq (1-\varepsilon) a_k\}$$

since $\varphi_n \leq \varphi_{n+1}$, $C_n^\varepsilon \subseteq C_{n+1}^\varepsilon$ and $\bigcup_{n=1}^\infty C_n^\varepsilon = E_k$, that means

$$\lim_{n\to\infty}\int_{E_k}\varphi_n = \lim_{n\to\infty}\int_{\mathbb{R}}\mathbf{1}_{E_k}\varphi_n \geq \lim_{n\to\infty}\int_{\mathbb{R}}\mathbf{1}_{C_n^\varepsilon}\varphi_n \geq \lim_{n\to\infty}(1-\varepsilon)a_k m(C_n^\varepsilon) = (1-\varepsilon)a_k m(E_k)$$

Since ε is chosen arbitrarily, so

$$\lim_{n\to\infty}\int_{E_k}\varphi_n \geq a_k m(E_k)$$

■

Corollary

Corollary 7. *For any $f \geq 0$ measurable, if $\forall n \geq 1, k = 0,1,2,\cdots,n2^n$ with $A_{n,k} = \left\{\frac{k}{2^n} \leq f < \frac{k+1}{2^n}\right\}$, then*

$$\int_{\mathbb{R}} f = \lim_{n\to\infty}\sum_{k=0}^{n2^n}\frac{k}{2^n}m(A_{n,k})$$

Now, we will introduce several properties of the integral of non-negative functions.

Proposition

Proposition 47. *(Well Defined) If $f,g \geq 0$ and are measurable such that $f = g$ for almost everywhere, then*

$$\int_{\mathbb{R}} f = \int_{\mathbb{R}} g$$

Proof. Let $\{\varphi_n : n \geq 1\}$, $\{\omega_n : n \geq 1\}$ be sequences of two simple functions such that φ_n, ω_n are increasing and $\lim_{n\to\infty}\varphi_n = f; \lim_{n\to\infty}\omega_n = g$. Define

$$h_n = \varphi_n\mathbf{1}_{f=g} + \omega_n\mathbf{1}_{f\neq g},$$

then we claim that h is again simple and increasing, also $\lim_{n\to\infty} h_n = g$, which implies

$$\int_{\mathbb{R}} g = \lim_{n\to\infty}\int_{\mathbb{R}} h_n = \lim_{n\to\infty}\left(\int_{f=g}\varphi_n + \int_{f\neq g}\omega_n\right),$$

as $n \to \infty$, we have

$$\lim_{n\to\infty}\int_{\mathbb{R}} h_n = \lim_{n\to\infty}\int_{\mathbb{R}}\varphi_n,$$

meaning that

$$\int_{\mathbb{R}} g = \lim_{n\to\infty}\int_{\mathbb{R}}\varphi_n$$

Meanwhile we have

$$\int_{\mathbb{R}} f = \lim_{n\to\infty}\varphi_n = \lim_{n\to\infty}\int_{f=g}\varphi_n$$

■

Proposition

Proposition 48. *(Linearity)*
Let $f, g \geq 0$ be measurable, and $a, b \in \mathbb{R}^+$, then

$$\int_{\mathbb{R}} af + bg = a\int_{\mathbb{R}} f + b\int_{\mathbb{R}} g$$

Proof. Take $\{\varphi_n\}, \{\omega\}$ as in the previous proof, and set $h_n = \{a\varphi_n + b\omega_n\}$, where h_n is simple and increasing, $\lim_{n\to\infty} h_n = af + bg$, then

$$\int_{\mathbb{R}} af + bg = \lim_{n\to\infty}\int_{\mathbb{R}} h_n = \lim_{a\to\infty}\left(a\int_{\mathbb{R}}\varphi_n + b\int_{\mathbb{R}}\omega_n\right) = a\int_{\mathbb{R}} f + b\int_{\mathbb{R}} g$$

■

Proposition

Proposition 49. *(Monotoncity)*
Let $f, g \geq 0$ be measurable with $f \leq g$ for almost everywhere, then

$$\int_{\mathbb{R}} f \leq \int_{\mathbb{R}} g$$

Proof. *W.L.O.G*, say $f \leq g$ point-wise, and let φ be a simple function, then $\{\varphi : \varphi \leq f\} \subseteq \{\varphi : \varphi \leq g\}$, where by definition we have

$$\int_{\mathbb{R}} f \leq \int_{\mathbb{R}} g$$

■

Proposition

Proposition 50. *(Markov's Inequality)*
Let $f \geq 0$ be measurable and $0 < a < +\infty$, then

$$m(\{f > a\}) \leq \frac{1}{a}\int_{\mathbb{R}} f$$

Proposition

Proposition 51. *Let $f \geq 0$ and is measurable, then we have :*

① $\int_{\mathbb{R}} f = 0$ *if and only if $f \equiv 0$ for almost everywhere.*

② *Let $A \in \mathscr{M}$, then $\int_{\mathbb{R}} f = 0$ implies whether $f \equiv 0$ for almost everywhere on A, or $m(A) = 0$.*

③ *If $\int_{\mathbb{R}} f < +\infty$, then f is finite for almost everywhere.*

①

Proof. $\Longleftarrow$ is trivial, now assume $\int_{\mathbb{R}} f = 0$, we want to show the set $A := \{f > 0\}$ has zero measure. For each n, the set $A_n := \{f \geq 1/n\}$ is increasing, and $\bigcup_{n=1}^{\infty} A_n = A$, by the properties of a measure, it also implies that $m(A) = \lim_{n\to\infty} m(A_n)$. Now assume $m(A) = \delta > 0$, then $\lim_{n\to\infty} m(A_n) = \delta > 0$, meaning that $\exists N$ large enough such that $m(A_N) > \delta/2$, since $f \geq f\mathbf{1}_{A_N} \geq \frac{1}{N}\mathbf{1}_{A_N}$, by monnotoncity,

$$\int_{\mathbb{R}} f \geq \int_{\mathbb{R}} \frac{1}{N}\mathbf{1}_{A_N} = \frac{1}{N} m(A_N) \geq \frac{1}{N}\frac{\delta}{2} > 0,$$

a contradiction. ■

②

Proof. By ①. we know that

$$\int_A f = 0 \Longleftrightarrow \mathbf{1}_A f = 0, \text{for almost everywhere}$$

then there are only two cases:

If $m(A) = 0$, then $\mathbf{1}_A f = 0$ holds for almost everywhere;

If $m(A) \neq 0$, then $\mathbf{1}_A = 1, \forall x \in A \implies f \equiv 0$ for almost everywhere ■

③

Proof. Let $A = \{f = +\infty\}$, assume $m(A) = \delta > 0$, then

$$\forall n \geq 1, f \geq f\mathbf{1}_A \implies \int_{\mathbb{R}} f \geq \int_{\mathbb{R}} n\mathbf{1}_A = nm(A) > 0,$$

a contradiction. ■

2.5.3 Integral for General Measurable (Integrable) Functions

Definition

Definition 20. *Let f be measurable, then define the integral of f as*

$$\int_{\mathbb{R}} f := \int_{\mathbb{R}} f^+ - \int_{\mathbb{R}} f^-$$

provided that at least one of $\int_{\mathbb{R}} f^+, \int_{\mathbb{R}} f^-$ is finite. Otherwise, we say the $\int_{\mathbb{R}} f$ is undefined.

Note that only having $\int_{\mathbb{R}} f$ defined is not sufficient for the desired properties (like linearity, monotoncity) to hold, so we give another definition:

Definition

Definition 21. *Let f be measurable, then f is integrable, denoted by $f \in \mathscr{L}^1(\mathbb{R})$, if*

$$\int_{\mathbb{R}} f^+ < +\infty \ ; \int_{\mathbb{R}} f^- < +\infty$$

Proposition

Proposition 52. *Two properties follow directly from the definition:*

Let f be measurable, then f is integrable if and only if $\int_{\mathbb{R}} |f| < +\infty$.

Let f be measurable, then f is integrable if and only if $\int_{\mathbb{R}} f$ is finite valued.

Now, we will further discuss of the properties of integrable functions $f \in \mathscr{L}^1(\mathbb{R})$.

Proposition

Proposition 53. *(Triangle-Inequality)*

$$\left| \int_{\mathbb{R}} f(x)dx \right| \leq \int_{\mathbb{R}} |f(x)|dx$$

Proposition

Proposition 54. *Let $f \in \mathscr{L}^1(\mathbb{R})$, then f is finite for almost everywhere.*

Proposition

Proposition 55. *(Linearity)*
Let $f, g \in \mathscr{L}^1(\mathbb{R})$, then $af + bg \in \mathscr{L}^1(\mathbb{R})$ and

$$\int_{\mathbb{R}} af + bg = a \int_{\mathbb{R}} f + b \int_{\mathbb{R}} g$$

Proof. First we will show that $af \in \mathscr{L}^1(\mathbb{R})$ and $\int_{\mathbb{R}} af = a \int_{\mathbb{R}} f$.

$$|af| \leq |a||f| \implies \int_{\mathbb{R}} |af| \leq \int_{\mathbb{R}} |a| \cdot |f| < +\infty$$

which means $af \in \mathscr{L}^1(\mathbb{R})$. Also note that we have $(af)^+ = \begin{cases} af^+, a \geq 0 \\ -af^-, a < 0 \end{cases}$ and same for $(af)^- =$

$\begin{cases} af^-, a \geq 0 \\ -af^+, a < 0 \end{cases}$, so we have

$$\int_{\mathbb{R}} af = \int_{\mathbb{R}} (af)^+ - \int_{\mathbb{R}} (af)^- = \begin{cases} \int_{\mathbb{R}} af^+ - af^-, a \geq 0 \\ \\ \int_{\mathbb{R}} -af^- - (-a)f^+, a < 0 \end{cases}$$

meaning that

$$\int_{\mathbb{R}} af = \begin{cases} a \int_{\mathbb{R}} f^+ - f^-, a \geq 0 \\ \\ -a \int_{\mathbb{R}} f^- - f^+, a < 0 \end{cases} = a \int_{\mathbb{R}} f$$

Similarly, $bg \in \mathscr{L}^1(\mathbb{R})$ and $\int_{\mathbb{R}} bg = b \int_{\mathbb{R}} g$. Now, $W.L.O.G$, assume $a = b = 1$, we will show that $f + g \in \mathscr{L}^1(\mathbb{R})$ and $\int_{\mathbb{R}} f + g = \int_{\mathbb{R}} f + \int_{\mathbb{R}} g$. Given $|f + g| \leq |f| + |g|$, by monotoncity we have $\int_{\mathbb{R}} |f + g| < +\infty$ and thus $f + g \in \mathscr{L}^1(\mathbb{R})$. Let $h = f + g$, then $|f|, |g|, |h| < +\infty$ for almost everywhere, same for the integrals. Then we have $h^+ - h^- = f^+ - f^- + g^+ - g^-$, meaning that $h^+ + f^- + g^- = f^+ + g^+ + g^-$, by linearity of integral of non-negative functions, we have

$$\int_{\mathbb{R}} h^+ + \int_{\mathbb{R}} f^- + \int_{\mathbb{R}} g^- = \int_{\mathbb{R}} f^+ + \int_{\mathbb{R}} g^+ + \int_{\mathbb{R}} h^-,$$

thus

$$\int_{\mathbb{R}} h^+ - h^- = \int_{\mathbb{R}} f^+ + g^+ - (f^- + g^-),$$

i.e

$$\int_{\mathbb{R}} h = \int_{\mathbb{R}} f + g$$

■

Proposition

Proposition 56. *Let $f \in \mathscr{L}^1(\mathbb{R})$, if $A \in \mathscr{M}, m(A) = 0$, then $\int_A f = 0$. In particular, if $f, g \in \mathscr{L}^1(\mathbb{R})$ and $f = g$ for almost everywhere, then $\int_{\mathbb{R}} f = \int_{\mathbb{R}} g$.*

Proposition

Proposition 57. *(Monotoncity)*
Let $f, g \in \mathscr{L}^1(\mathbb{R})$, if $f \leq g$ for almost everywhere, then

$$\int_{\mathbb{R}} f \leq \int_{\mathbb{R}} g$$

Proof. Set $h = g - f$, then $h \in \mathscr{L}^1(\mathbb{R})$ and $h \geq 0$ for almost everywhere, then

$$\int_{\mathbb{R}} h = \int_{\mathbb{R}} g - f \geq 0 \implies \int_{\mathbb{R}} g \geq \int_{\mathbb{R}} f$$

∎

Convergence Theorems of Integral

In this section, we will discuss under what condition is $\int_{\mathbb{R}} \lim f_n = \lim \int_{\mathbb{R}} f_n$ true. Also, in this part $\lim f_n = f$ pointwise can all be replaced with ”almost everywhere”.

Theorem

Theorem 20. *(Monotone Convergence Theorem (MON))*
Let $\{f_n : n \geq 1\}, f$ be non-negative and measurable functions, if f_n is increasing and $\lim_{n\to\infty} f_n =$ *f, then*

$$\int_{\mathbb{R}} f = \lim_{n\to\infty} \int_{\mathbb{R}} f_n$$

Proof. By monotoncity of non-negative functions, we know that $\lim_{n\to\infty} \int_{\mathbb{R}} f_n$ exists and $\lim_{n\to\infty} \int_{\mathbb{R}} f_n \leq \int_{\mathbb{R}} f$ for almost everywhere, now we will show the reverse inequality. Now for each $n \geq 1$, select $\{\varphi_{n,k} : k \geq 1\}$ to be a sequence of simple functions such that $\varphi_{n,k}$ is increasing and $\lim_{k\to\infty} \varphi_{n,k} = f_n$. Define $g_k := \max\{\varphi_{1,k}, \varphi_{2,k}, \cdots, \varphi_{k,k}\}$, where we know g_n is simple and increasing, also $g_k \leq f$, thus $\lim_{n\to\infty} g_k \leq f$ and thus for every fixed $n \geq 1$, we have

$$\lim_{k\to\infty} g_k \geq \lim_{k\to\infty} \varphi_{n,k} = f_n,$$

and thus

$$\lim_{k\to\infty} g_k \geq \lim_{k\to\infty} f_n = f$$

and we have $\lim_{k\to\infty} g_k = f$, where $\lim_{k\to\infty} \int_{\mathbb{R}} g_k = \int_{\mathbb{R}} f$, since $\varphi_{1,k}, \varphi_{2,k}, \cdots \leq f_k$, it means that $g_k \leq f_k$ and thus $\int_{\mathbb{R}} g_k \leq \int_{\mathbb{R}} f_k$, where

$$\lim_{k\to\infty} \int_{\mathbb{R}} g_k = \int_{\mathbb{R}} f \leq \lim_{k\to\infty} \int_{\mathbb{R}} f_k$$

∎

Corollary

Corollary 8. *Let* $\{f_n : n \geq 1\}, f$ *be measurable functions such that* f_n *is increasing and* $\lim_{n\to\infty} f_n = f$ *and* $\int_{\mathbb{R}} f_1^- < +\infty$, *then*

$$\int_{\mathbb{R}} f = \lim_{n\to\infty} \int_{\mathbb{R}} f_n$$

Proof. Since f_n is increasing, in particular $f \geq f_1$ and $f^- \leq f_1^-$, and all of which are finite valued for almost everywhere, thus $\int_{\mathbb{R}} f_n^- \leq \int_{\mathbb{R}} f_1^- < +\infty$. For all $n \geq 1$, set $\widetilde{f_n} := f_n + f_1^-$, and $\widetilde{f_n} \geq 0$ and is increasing, also $\lim_{n\to\infty} \widetilde{f_n} = f + f_1^- := \widetilde{f}$, now by MON, we have

$$\int_{\mathbb{R}} \widetilde{f} = \lim_{n\to+\infty} \int_{\mathbb{R}} \widetilde{f_n},$$

thus

$$\int_{\mathbb{R}} (f + f_1^-) = \lim_{n\to\infty} \int_{\mathbb{R}} (f_n + f_1^-)$$

If we could apply linearity, then we are already done, but that's not the case. We will do the following instead:

Let $\widetilde{f_n} = f_n + f_1^- = f_n^+ + f_n^- + f_1^-$, where $\widetilde{f_n} + f_n^- = f_n^+ + f_1^-$, now by applying linearity, we get

$$\int_{\mathbb{R}} \widetilde{f_n} = \int_{\mathbb{R}} f_n^+ - \int_{\mathbb{R}} f_n^- + \int_{\mathbb{R}} f_1^-,$$

similarly,

$$\int_{\mathbb{R}} \widetilde{f} = \int_{\mathbb{R}} + \int_{\mathbb{R}} f_1^-$$

■

Here is a remark: The condition $\int_{\mathbb{R}} f_1^- < +\infty$ is necessary. We may consider $f_n(x) = \begin{cases} -1, x \geq n \\ 0, x < n \end{cases}$, where f_n is increasing and $\lim_{n\to\infty} f_n = 0$, however $\int_{\mathbb{R}} f_n = -\infty$, where MON fails.

Theorem

Theorem 21. *(Reverse MON)*
Let $\{f_n : n \geq 1\}, f$ *be measurable functions,* f_n *is decreasing and* $\lim_{n\to\infty} f_n = f$, *if* $\int_{\mathbb{R}} f_1^+ < +\infty$, *then*

$$\int_{\mathbb{R}} f = \lim_{n\to\infty} \int_{\mathbb{R}} f_n$$

Proof. This proof is left as an exercise. ■

Theorem

Theorem 22. *(Fatou's)*
Let $\{f_n : n \geq 1\}$ be a sequence of non-negative and measurable functions, then

$$\int_{\mathbb{R}} \liminf_n (f_n) \leq \liminf_n \left(\int_{\mathbb{R}} f_n \right)$$

Proof. $\forall m \geq 1$, let $g_m := \inf_{n \geq m} f_n$, where $g_m \geq 0$ and is increasing. Also $\lim_{m \to \infty} g_m = \liminf_n (f_n)$, by MON, we have

$$\int_{\mathbb{R}} \liminf_n (f_n) = \lim_{m \to \infty} \int_{\mathbb{R}} g_m$$

$\forall n \geq m$, $g_m \leq f_n$, thus by monotoncity of non-negative functions,

$$\int_{\mathbb{R}} g_m \leq \int_{\mathbb{R}} f_n, \forall n \geq m,$$

which implies

$$\int_{\mathbb{R}} g_m \leq \inf_{n \geq m} \int_{\mathbb{R}} f_n,$$

i.e

$$\int_{\mathbb{R}} \liminf_n (f_n) \leq \liminf_n \left(\int_{\mathbb{R}} f_n \right)$$

■

Corollary

Corollary 9. *Let $\{f_n : n \geq 1\}$ be a sequence of measurable functions, if there exists a measurable function g, such that $\int_{\mathbb{R}} g^- < +\infty$, $\forall n \geq 1 : f_n \geq g$. Then*

$$\int_{\mathbb{R}} \liminf_n (f_n) \leq \liminf_n \left(\int_{\mathbb{R}} f_n \right)$$

Proof. Since $f \geq g$ for all $n \geq 1$, we then have $f_n^- \leq g^-, f_n^- < +\infty, \int_{\mathbb{R}} f_n^- < +\infty$. Set $\widetilde{f_n} := f_n + g^-$, then $\widetilde{f_n} \geq 0$. By $Fatou's$ theorem,

$$\int_{\mathbb{R}} \liminf_n (\widetilde{f_n}) \leq \liminf_n \left(\int_{\mathbb{R}} \widetilde{f_n} \right),$$

then check linearity of

$$\int_{\mathbb{R}} \liminf_n (f_n + g^-) \leq \liminf_n \left(\int_{\mathbb{R}} (f_n + g^-) \right)$$

■

Theorem

Theorem 23. *(Reverse Fatou)*
Let $\{f_n : n \geq 1\}$ *be measurable, if there exists a measurable function* g *such that* $\int_{\mathbb{R}} g^+ < +\infty$*, and* $\forall n \geq 1 : f_n \leq g$*, then*

$$\int_{\mathbb{R}} \limsup_n (f_n) \geq \limsup_n \left(\int_{\mathbb{R}} f_n \right)$$

Proof. This proof is left as an exercise. ■

Theorem

Theorem 24. *(Dominated Convergence Theorem)*
Let $\{f_n : n \geq 1\}, f$ *be measurable functions, if there exists* $g \in \mathscr{L}^1(\mathbb{R})$ *such that* $\forall n \geq 1 : |f_n| \leq |g|$ *and* $f_n \longrightarrow f$ *for almost everywhere, then* $f_n \longrightarrow f$ *in* $\mathscr{L}^1(\mathbb{R})$*, i.e* $\lim\limits_{n\to\infty} \int_{\mathbb{R}} |f_n - f| = 0$*, in particular*

$$\int_{\mathbb{R}} f = \lim_{n\to\infty} \int_{\mathbb{R}} f_n$$

Proof. Since $|f_n| \leq |g|$ and $f = \lim_{n\to\infty} f_n$, then $|f| \leq |g|$ for almost everywhere, meaning $\int_{\mathbb{R}} |f_n| \leq \int_{\mathbb{R}} |g|$ for almost everywhere and $\int_{\mathbb{R}} |f| \leq \int_{\mathbb{R}} |g|$ for almost everywhere, and thus $f_n, f \in \mathscr{L}^1(\mathbb{R})$. Notice that $|f_n - f| \leq 2|g|$, thus by applying reverse *Fatou* to $\{|f_n - f|\}$, we have

$$\int_{\mathbb{R}} \limsup_n (|f_n - f)| \geq \limsup_n \left(\int_{\mathbb{R}} |f_n - f| \right)$$

Which means $\limsup\limits_n \int_{\mathbb{R}} |f_n - f| = 0$, and

$$\left| \int_{\mathbb{R}} f_n - \int_{\mathbb{R}} f \right| = \left| \int_{\mathbb{R}} (f_n - f) \right| = \int_{\mathbb{R}} |f_n - f| \longrightarrow_{n\to\infty} 0,$$

meaning that $\lim\limits_{n\to\infty} \int_{\mathbb{R}} f_n = \int_{\mathbb{R}} f$. ■

Here is a remark: In order to apply *DOM*, we must find $g \in \mathscr{L}^1(\mathbb{R})$ such that $|g| \geq |f_n|$ for all $n \geq 1$, take $f_n = \mathbf{1}_{[n,2n]}$, then $\lim_{n\to\infty} f_n = 0$, but $\int_{\mathbb{R}} f_n = n$, and *DOM* does not apply.

Proposition

Proposition 58. *Let* $f \in \mathscr{L}^1(\mathbb{R})$*, and* $\{h_n : n \geq 1\}$ *is a sequence of measurable functions that are uniformly bounded, i.e* $\exists M > 0$*, such that* $\forall n \geq 1 : |h_n| \leq M$ *for almost everywhere. If* $h_n \longrightarrow h$ *for almost everywhere for some measurable functions* h*, then*

$$\lim_{n\to\infty} \int_{\mathbb{R}} (f \cdot h_n) = \int_{\mathbb{R}} (f \cdot h)$$

Proof. $\forall n \geq 1$, $|fh_n| \leq M|f| \in \mathscr{L}^1(\mathbb{R})$, the conclusion follows from *DOM*. ■

Corollary

Corollary 10. *If $f \in \mathscr{L}^1(\mathbb{R})$, then $\forall \varepsilon > 0$, $\exists$ compact set $K \subseteq \mathbb{R}$ such that*

$$\int_{K^C} |f| < \varepsilon$$

Corollary

Corollary 11. *If $f \in \mathscr{L}^1(\mathbb{R})$, then $\forall \varepsilon > 0$, $\exists N \geq 1$, such that*

$$\int_{|f|>N} |f| < \varepsilon$$

Corollary

Corollary 12. *If $\{A_n\} \subseteq \mathscr{M}$, such that $\{A_n\}$ is increasing, then*

$$\int_{\bigcup_{n=1}^{\infty} A_n} f = \lim_{n\to\infty} \int_{A_n} f$$

Similarly, if $\{A_n\}$ is decreasing, then

$$\int_{\bigcap_{n=1}^{\infty} A_n} f = \lim_{n\to\infty} \int_{A_n} f$$

Corollary

Corollary 13. *If $\{B_n\} \subseteq \mathbb{M}$ and are pairwise disjoint, then*

$$\int_{\bigcup_{n=1}^{\infty} B_n} f = \sum_{n=1}^{\infty} \int_{B_n}^{n} f$$

Proposition

Proposition 59. *Assume f is a non-negative, measurable and finite valued function a.e, $\forall k \in \mathbb{Z}$, set $A_k := \{x \in \mathbb{R} : 2^k \leq f(x) < 2^{k+1}\}$ where $A_k \in \mathscr{M}$, then $\int_{\mathbb{R}} f < +\infty$ if and only if $\sum_{k\in\mathbb{Z}} 2^k m(A_k) < +\infty$.*

Proof. $(\Longrightarrow)$: Assume that $\int_{\mathbb{R}} f$ converge, note that A_k's are disjoint, and $\bigcup_{k\in\mathbb{Z}} A_k = \{0 < f < +\infty\}$, then

$$\int_{\mathbb{R}} f = \int_{\{0<f<+\infty\}} f = \sum_{k\in\mathbb{Z}} \int_{A_k} f$$

For each k, $\forall x \in A_k$, $2^k \leq f(x) < 2^{k+1}$, and

$$2^k m(A_k) \leq \int_{A_k} f < 2^{k+1} m(A_k),$$

thus

$$\sum_{k\in\mathbb{Z}} 2^k m(A_k) \leq \sum_{k\in\mathbb{Z}} \int_{A_k} f(x)dx < +\infty.$$

$(\Longleftarrow)$: Assume that $\sum_{k\in\mathbb{Z}} 2^{k+1} m(A_k) < +\infty$, then

$$\int_{\mathbb{R}} f = \int_{\{0<f<+\infty\}} f = \sum_{k\in\mathbb{Z}} \int_{A_k} f < \sum_{k\in\mathbb{Z}} 2^{k+1} m(A_k) = 2\sum_{k\in\mathbb{Z}} 2^k m(A_k) < +\infty$$

■

This proposition is really useful in dealing some integrals, we shall illustrate this by some examples:

Example 1. Consider the function defined by $f(x) = |x|^{-\alpha} \cdot \mathbf{1}_{[-1,1]}(x)$, $\alpha > 0$, where f is finite valued. What can we say about the convergence of $\int_{\mathbb{R}} f$ under Lebesgue integral?

Solution: For any $k \in \mathbb{Z}$, we define $A_k := \{2^k \leq f < 2^{k+1}\} = \{2^k < |x|^{-\alpha} < 2^{k+1}\}$ given that $|f| \geq 1, x \in [-1,1]$. By solving this we get

$$A_k := [-2^{-k/\alpha}, -2^{-(k+1)/\alpha}) \cup (-2^{-(k+1)/\alpha}, 2^{-k/\alpha}], k \geq 0,$$

and $A_k \varnothing$ if $k < 0$, this implies that

$$\begin{aligned}\sum_{k\in\mathbb{Z}} 2^k m(A_k) &= \sum_{k=0}^{\infty} 2^k \cdot 2(1-2^{-1/\alpha})2^{-k/\alpha} \\ &= 2(1-2^{-1/\alpha}) \sum_{k=0}^{\infty} 2^{k(1-(1/\alpha)}\end{aligned}$$

converge if and only if $0 < \alpha < 1$, and $\int_{[-1,1]} |x|^{-\alpha} dx$ converges if and only if $0 < \alpha < 1$.

Example 2. Consider the series of functions given by $f_n(x) = \left(1 + \frac{x}{n}\right)^{-n} \cdot \sin\left(\frac{x}{n}\right)$, what do we know about the convergence of $\lim_{n\to\infty} \int_{(0,+\infty)} f_n(x)dx$ under Lebesgue integral?

Solution: We observe that

$$f(x) \leq \left(1 + \frac{x}{n}\right)^{-n} \leq \left(1 + \frac{x}{2}\right)^{-2}, \forall x > 0, n \geq 2.$$

Let $g(x) = \left(1 + \frac{x}{2}\right)^{-2}$, we claim that $g(x) \in \mathscr{L}^1((0,+\infty))$, then by *DOM*, we get

$$\lim_{n\to\infty} \int_{(0,+\infty)} f_n = \int_{(0,+\infty)} \lim_{n\to\infty} f_n = 0,$$

thus it converges.

Example 3. Consider the series of functions given by $f_n(x) = x^{-c} \cdot (\cosh x)^{-1/n}$, $c > 0$, what do we know about the convergence of $\lim_{n\to\infty} \int_{(1,+\infty)} f_n(x)$?

Solution: We know that $\forall x \geq 1$, $\cosh x \geq 1$, and the sequence $(\cosh x)^{-1/n}$ is increasing, with $\lim_{n\to\infty} (\cosh x)^{-1/n} = 1$, thus $\lim_{n\to\infty} f_n(x) = x^{-c}, \forall x \geq 1$. Let $g(x) = x^{-c}$, consider two cases:

• When $c > 1$, we know that $g(x) \in \mathscr{L}^1((1,+\infty))$, by *DOM* we have

$$\lim_{n\to\infty} \int_{(1,+\infty)} f_n(x)dx = \int_{(1,+\infty)} \lim_{n\to\infty} f_n(x)dx = \int_{(1,+\infty)} x^{-c}dx < +\infty$$

and it converges.

• when $0 < c \leq 1$, by *Fatou's* lemma, we have

$$\liminf_n \int_{(1,+\infty)} f_n \geq \int_{(1,+\infty)} \liminf_n f_n = \int_{(1,+\infty)} x^{-c}dx = +\infty$$

and it diverges.

We may see that the conditions for certain integral to converge are the same for both Riemann integral and Lebesgue integral. What are the main differences and similarities between the two?

Riemann & Lebesgue Integral

Let's recall some properties from the Riemann Integral:

Proposition

Proposition 60. *Let f be bounded on $[a,b]$, then f is Riemann integrable on $[a,b]$ if f is continuous on $[a,b]$, or f is monotone on $[a,b]$, or f is continuous except at possibility finitely many points in $[a,b]$.*

Recall that the function $f = \mathbf{1}_{\mathbb{Q}\cap[0,1]}$ is not Riemann integrable, but f is Lebesgue integrable, since $|f| \leq \mathbf{1}_{[0,1]} \in \mathscr{L}^1([0,1])$. Here are three remarks:

1. There exists a bounded function on a bounded interval, such that it is not Riemann integrable.

2. In general, g is Riemann integrable and $|f| \leq |g|$ does not apply f is Riemann integrable (like the example we constructed).

3. In general, theorems like DOM, MON do not apply to Riemann integral. Here is an example: Take $\{q_n\}$ to be an enumeration of $\mathbb{Q}\cap[0,1]$, define

$$f_n(x) = \begin{cases} 1 : x \in \{q_1, q_2, \cdots, q_n\} \\ 0 : \text{otherwise} \end{cases},$$

in this case $f_n(x)$ is increasing and $\lim_{n\to\infty} f_n = f = \mathbf{1}_{\mathbb{Q}\cap[0,1]}$, now $\forall n \geq 1$, f_n is only discontinuous at finitely many points, so f_n is Riemann integrable on $[0,1]$ and $\int_0^1 f_n(x)dx = 0$. However $\lim_{n\to\infty} f_n(x)$ is NOT Riemann integrable.

Theorem

Theorem 25. *Assume f is bounded and Riemann integrable on $[a,b]$, then f is Lebesgue integrable on $[a,b]$, and moreover both integral match in value.*

Proof. Say f is Riemann integrable on $[a,b]$, then $\exists M$ such that $|f| < M$ on $[a,b]$, and there are step functions $\{\varphi_n\}, \{\omega_n\}$ such that

$$\varphi_n \leq f \leq \omega_n$$

on $[a,b]$, $W.L.O.G$, $|\varphi_n|, |\omega_n| \leq M, \forall n \geq 1$, and

$$\lim_{n\to\infty} \int_a^b \varphi_n = \int_a^b f = \lim_{n\to\infty} \int_a^b \omega_n$$

If we denote $\varphi = \lim_{n\to\infty} \int_a^b \varphi_n; \omega = \lim_{n\to\infty} \int_a^b \omega_n$, then since $\{\varphi_n\}, \{\omega_n\}$ are simple functions, thus they are measurable, and it also applies to φ, ω, where $\varphi \leq f \leq \omega$. Observe that Lebesgue integral and Riemann integral coincide on step functions, thus

$$\int_a^b \varphi_n = \int_{[a,b]} \varphi_n; \int_a^b \omega_n = \int_{[a,b]} \omega_n,$$

thus by DOM,

$$\begin{aligned}
\int_{[a,b]} \varphi = \lim_{n\to\infty} \int_{[a,b]} \varphi_n &= \lim_{n\to\infty} \int_a^b \varphi_n \\
&= \int_a^b f = \lim_{n\to\infty} \int_a^b \omega_n \\
&= \lim_{n\to\infty} \int_{[a,b]} \omega_n \\
&= \int_{[a,b]} \omega_n
\end{aligned}$$

Note that $\varphi \leq \omega \implies \int_{a,b]} (\omega - \varphi) \geq 0$, also $\int_{[a.b]} (\omega - \varphi) = 0 \implies \omega = \varphi$ a.e on $[a,b]$, where $f = \omega = \varphi$ a.e on $[a,b]$, f is measurable. Since $|f| \leq M$ on $[a,b]$, thus f is Lebesgue integrable, and it has the same value as the Riemann integral. ∎

Theorem 25 is useful in many ways. Consider the following example:

Example: Determine $\lim_{n\to\infty}\int_{[0,+\infty)}\frac{n}{1+n^2x^2}dx$.

Solution: Observe that

$$\int_{[0,+\infty)}\frac{n}{1+n^2x^2}dx=\int_{[0,T]}\frac{n}{1+n^2x^2}dx+\int_{[T,+\infty)}\frac{n}{1+n^2x^2}dx$$

As $T\to\infty$, by DOM we have $\int_{[T,+\infty)}\frac{n}{1+n^2x^2}dx\to 0$, as $n\to+\infty$, and

$$\int_{[0,T]}\frac{n}{1+n^2x^2}dx=\int_0^T\frac{ndx}{1+n^2x^2}=\arctan(nT)$$

By setting $n\to\infty$, we have $\int_{[0,T]}\frac{n}{1+n^2x^2}dx=\frac{\pi}{2}$, and

$$\lim_{n\to\infty}\int_{[0,+\infty)}\frac{n}{1+n^2x^2}dx=\frac{\pi}{2}.$$

$\mathscr{L}^p$ Space

Definition

Definition 22. *Assume f is a measurable function, and $1\le p\le+\infty$. Then we say that f is p-integrable, written as $f\in\mathscr{L}^p(\mathbb{R})$, if*

$$\int_{\mathbb{R}}|f|^p<+\infty.$$

i.e $|f|^p\in\mathscr{L}^1(\mathbb{R})$, and we define the $\mathscr{L}^p$ norm of f, given by

$$\|f\|_p=\left(\int_{\mathbb{R}}|f|^p\right)^{\frac{1}{p}}.$$

When $p=1$, we just have $\|f\|_1=\int_{\mathbb{R}}f$, and this is a norm on $\mathscr{L}^1(\mathbb{R})$.

Definition

Definition 23. *Say $\|\cdot\|_p$ is a norm on $\mathscr{L}^p(\mathbb{R})$, if $f,g \in \mathscr{L}^p(\mathbb{R})$, we have the following:*

① $\|f\|_p = 0$, *then* $f \equiv 0$ *a.e;*

② $\forall a \in \mathbb{R}$, $\|af\|_p = |a| \cdot \|f\|_p$;

③ *(Triangle Inequality)* $\|f+g\|_p \leq \|f\|_p + \|g\|_p$.

In fact, for any $1 \leq p \leq \infty$, $\|\cdot\|_p$ is indeed a norm defined on $\mathscr{L}^p$, when $p = \infty$, we define

$$\|f\|_\infty := \inf\{a \in \overline{\mathbb{R}} : |f| \leq a \;\; a.e\}$$

We will then study some important inequalities.

Theorem

Theorem 26. *(Hölder's Inequality)*
Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, where q is called the Hölder conjugate. If $f \in \mathscr{L}^p(\mathbb{R}), g \in \mathscr{L}^q(\mathbb{R})$, we then have

$$\|f \cdot g\|_1 \leq \|f\|_p \cdot \|g\|_q$$

Remark : If $p = q = 2$, then it is just Cauchy-Schwarz Inequality.
Before the proof of Hölder's Inequality, we need some results from other famous inequalities:

Theorem

Theorem 27. *(Young's Inequality)*
$\forall a,b \geq 0$, if $p,q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, then

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

Theorem

Theorem 28. *(Jensen's Inequality)*
Let f defined on $I \subseteq \mathbb{R}$, then $\forall x_1, x_2, \cdots, x_n \in I$, $a_1, a_2, \cdots, a_n > 0, a_1 + a_2 + \cdots + a_n = 1$.

- *If f is convex, then $f(a_1x_1 + a_2x_2 + \cdots + a_nx_n) \leq a_1f(x_1) + a_2f(x_2) + \cdots + a_nf(x_n)$;*

- *If f is concave, then $f(a_1x_1 + a_2x_2 + \cdots + a_nx_n) \geq a_1f(x_1) + a_2f(x_2) + \cdots + a_nf(x_n)$;*

We will not prove theorem 27 & 28, however let's give Hölder's Inequality a proof:

Proof. Since $f \in \mathscr{L}^p, g \in \mathscr{L}^q$, then set $\widetilde{f} = \dfrac{f}{\|f\|_p}, \widetilde{g} = \dfrac{g}{\|g\|}$, then for almost everywhere. By Young's inequality, we have

$$|\widetilde{f}\cdot\widetilde{g}| \leq \frac{|\widetilde{f}|^p}{p} + \frac{|\widetilde{g}|^q}{q}, \text{where}$$

$$\frac{\int_{\mathbb{R}}|f\cdot g|}{\|f\|_p\cdot\|g\|_q} \leq \frac{1}{p}\frac{\int_{\mathbb{R}}|f|^p}{\|f\|_p^p} + \frac{1}{q}\frac{\int_{\mathbb{R}}|g|^q}{\|g\|_q^q} = \frac{1}{p}+\frac{1}{q} = 1,$$

and thus we get Hölder's Inequality.

■

Remark: Hölder's Inequality also holds for $p = 1, q = \infty$. The proof is a bit challenging, those who interested are urged to try.

Theorem

Theorem 29. *(Minkouski's Inequality)*
Let $1 \leq p < +\infty$, $f, g \in \mathscr{L}^p(\mathbb{R})$, then $f+g \in \mathscr{L}^p(\mathbb{R})$ and

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p$$

Proof. First we notice that $|f+g|^p \leq 2^p(|f|+|g|^p)$, meaning that $f+g \in \mathscr{L}^p(\mathbb{R})$, thus

$$\begin{aligned}
\int_{\mathbb{R}}|f+g|^p &= \int_{\mathbb{R}}|f+g|\cdot|f+g|^{p-1} \\
&\leq \int_{\mathbb{R}}|f|\cdot|f+g|^{p-1} + \int_{\mathbb{R}}|g|\cdot|f+g|^{p-1} \\
&\text{By Hölder's Inequality,} \\
&\leq \left(\int_{\mathbb{R}}|f|^p\right)^{\frac{1}{p}}\left(\int_{\mathbb{R}}|f+g|^{(p-1)q}\right)^{\frac{1}{q}} + \left(\int_{\mathbb{R}}|g|^p\right)^{\frac{1}{p}}\left(\int_{\mathbb{R}}|f+g|^{(p-1)q}\right)^{\frac{1}{q}} \\
&= (\|f\|_p+\|g\|_p)\left(\int_{\mathbb{R}}|f+g|^p\right)^{\frac{1}{q}}
\end{aligned}$$

Thus it implies Minkouski's inequality.

■

Remark: Minkouski's inequality also holds for $p = +\infty$.

Theorem

Theorem 30. *Let $1 \leq p < +\infty$, if a sequence $\{g_k\} \subseteq \mathscr{L}^p(\mathbb{R})$ such that $\sum_k \|g_k\|_p$ converges, then $\exists G \in \mathscr{L}^p(\mathbb{R})$ such that as $m \to \infty$, $G_m := \sum_{k=1}^m g_k \to G$ a.e and in $\mathscr{L}^p(\mathbb{R})$.*

Proof. Set $\widetilde{G_m} := \sum_{k=1}^m |g_k|$, and $\widetilde{G} := \sum_{k=1}^\infty |g_k|$, then $\widetilde{G_m}$ is increasing and $\lim_{m\to\infty}\widetilde{G_m} = \widetilde{G}$, thus by

MON, we have

$$\begin{aligned}\int_{\mathbb{R}} (\widetilde{G})^p &:= \lim_{m\to\infty} \int_{\mathbb{R}} (\widetilde{G_m})^p \\ &= \lim_{m\to\infty} (\|\widetilde{G_m}\|_p)^p \\ &\le \lim_{m\to\infty} \left(\sum_{k=1}^{m} \|g_k\|_p\right)^p = \left(\sum_{k=1}^{\infty} \|g_k\|^p\right)^p < +\infty\end{aligned}$$

Thus $\widetilde{G} \in \mathscr{L}^p(\mathbb{R})$ and $\|\widetilde{G}\|_p \le \sum_{k=1}^{\infty} \|g_k\|_p$, $\widetilde{G}$ is finite valued a.e, then $\sum_{k=1}^{\infty} g_k$ is absolutely convergent (a.e).

Now set $G = \lim_{n\to\infty} G_m = \sum_{k=1}^{\infty} g_k$ a.e, moreover we know that $|G| \le \widetilde{G} \implies G \in \mathscr{L}^p(\mathbb{R})$, now $|G - G_m| \le \sum_{k=m+1}^{\infty} |g_k|$, and $\forall \varepsilon > 0$, there exists M large enough, such that

$$\sum_{k=M+1}^{\infty} \|g_k\|_p < \varepsilon,$$

which implies

$$\begin{aligned}\int_{\mathbb{R}} |G - G_M|^p \le \int_{\mathbb{R}} \left(\sum_{k=M+1}^{\infty} |g_k|\right)^p &= \lim_{L\to\infty} \int_{\mathbb{R}} \left(\sum_{k=M+1}^{L} |g_k|\right)^p \\ &\le \lim_{L\to\infty} \left(\sum_{k=M+1}^{L} \|g_k\|_p\right)^{\frac{1}{p}} = \left(\sum_{k=M+1}^{\infty} \|g_k\|_p\right)^{\frac{1}{p}} \le \varepsilon\end{aligned}$$

Which implies $\|G - G_m\|_p \le \sum_{k=M+1}^{\infty} \|g_k\|_p \le \varepsilon$, thus $G_m \to G$ in $\mathscr{L}^p(\mathbb{R})$. ■

Now we give the definition of a complete $\mathscr{L}^p(\mathbb{R})$ space:

Definition

Definition 24. *The space $\mathscr{L}^p$ is a complete norm space under $\|\cdot\|_p$, if a sequence $\{f_n\} \subseteq \mathscr{L}^p(\mathbb{R})$ is Cauchy under $\|\cdot\|_p$, then $\exists f \in \mathscr{L}^p(\mathbb{R})$ such that $f_n \to f$ in $\mathscr{L}^p(\mathbb{R})$ and $\lim_{n\to\infty} \|f_n - f\|_p = 0$.*

Theorem

Theorem 31. *Let $1 \le p < +\infty$, then $\mathscr{L}^p(\mathbb{R})$ is a complete norm space under $\|\cdot\|_p$.*

Proof. Let $\{f_n\} \subseteq \mathscr{L}^p(\mathbb{R})$ be Cauchy under $\|\cdot\|_p$, then we may choose a sub-sequence $\{n_k\}$ such that

$$\forall k \ge 1 : \|f_{n_{k+1}} - f_{n_k}\|_p \le 2^{-k}$$

Now set $g_k := f_{n_{k+1}} - f_{n_k}$, $G_m = \sum_{k=1}^{m} g_k$, then by theorem 30, there exists G, such that $G_m \to G$ a.e and in $\mathscr{L}^p(\mathbb{R})$. In fact,

$$G_m = \sum_{k=1}^{m} (f_{n_{k+1}} - f_{n_k}) = f_{n_{m+1}} - f_{n_1}$$

That means $G = \lim_{m\to\infty} G_m = \lim_{m\to\infty} f_{n_{m+1}} - f_{n_1}$. Define $f = G + f_{n_1}$, then $f = \lim_{m\to\infty} f_{n_m}$ a.e and since $G_m \to G$ in $\mathscr{L}^p(\mathbb{R})$, so $f_{n_m} \to f$ in $\mathscr{L}^p(\mathbb{R})$. Now $\forall \varepsilon > 0$, there exists N large enough such that

$$\sup_{k,l\geq n} \|f_k - f_l\|_p < \varepsilon$$

Choose m large enough as well, such that $n_m > n$, then $\|f_{n_m} - f\|_p < \varepsilon$, it means

$$\|f_n - f\|_p \leq \|f_n - f_{n_m}\|_p + \|f_{n_m} - f\|_p \leq 2\varepsilon$$

■

Remark: $\mathscr{L}^\infty$ is also a complete normed space under $\|\cdot\|_\infty$.

Next we want to identify some dense subspaces of $\mathscr{L}^p(\mathbb{R})$, where $1 \leq p < +\infty$.

Corollary

Corollary 14. *Bounded and compactly supported functions are dense in $\mathscr{L}^p(\mathbb{R})$.*

Proof. To prove this, we will show that $\forall f \in \mathscr{L}^p(\mathbb{R})$, there exists a sequence of bounded and compact supported functions that converge to f. For each n, set $f_n = \mathbf{1}_{[-n,n]}(x) \cdot f(x) \cdot \mathbf{1}_{|f|\leq n}(x)$, and f_n is bounded and compactly supported on $[-n,n]$, and we claim that $f_n \to f$ in $\mathscr{L}^p(\mathbb{R})$, since

$$\int_{\mathbb{R}} |f_n - f|^p \leq \int_{\mathbb{R}\setminus[-n,n]} |f|^p + \int_{|f|>n} |f|^p \to 0 \quad \text{as } n \to \infty.$$

■

Corollary

Corollary 15. *Simple functions are dense in $\mathscr{L}^p(\mathbb{R}), 1 \leq p < +\infty$.*

Proof. Let f_n be the sequence constructed the same as above. Then for each $n \geq 1, k = 0, 1, \cdots, n2^n - 1$, set

$$A_{n,k} = \left\{ x \in [-n,n] : \frac{k}{2^n} \leq f_n^+ < \frac{k+1}{2^n} \right\}$$

and

$$B_{n,k} = \left\{ x \in [-n,n] : \frac{k}{2^n} \leq f_n^- < \frac{k+1}{2^n} \right\}$$

We define

$$\varphi_n^+ = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbf{1}_{A_{n,k}} \quad \varphi_n^- = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbf{1}_{B_{n,k}}$$

and we say that $\varphi_n = \varphi_n^+ - \varphi_n^-$ is again simple, $|\varphi_n| \leq n$, and thus $\varphi_n \in \mathscr{L}^p(\mathbb{R})$, and $\varphi_n \to f$, $\forall x \in \mathbb{R}$, $\lim_{n\to\infty} \varphi_n(x) = f(x)$, also

$$|f_n(x) - \varphi_n(x)| \leq |f_n^+(x) - \varphi_n^+(x)| + |f_n^-(x) - \varphi_n^-(x)| \leq 2 \cdot 2^{-n}$$

which implies

$$\begin{aligned}
\|f-\varphi_n\|_p \leq \underbrace{\|f-f_n\|_p}_{=0,\text{ as } n\to\infty} + \|f_n-\varphi_n\|_p & \\
&\leq \left(\int_{\mathbb{R}} |f_n-\varphi_n|^p\right)^{\frac{1}{p}} \\
&\leq \left((2\cdot 2^{-n})^p \cdot m([-n,n])\right)^{\frac{1}{p}} \\
&= (2^p\cdot 2^{-np}\cdot 2^n)^{\frac{1}{p}} \to 0, \text{as } n\to\infty.
\end{aligned}$$

■

Theorem

Theorem 32. *Denote by $C_C(\mathbb{R})$, the space of continuous and compactly supported functions, then $C_C(\mathbb{R})$ is dense in $\mathscr{L}^p(\mathbb{R})$, $1\leq p<+\infty$.*

Proof. Given $f\in\mathscr{L}^p(\mathbb{R})$, choose a sequence $\{\varphi_n\}$ the same as above, recall that $\forall n\geq 1$, there exists a step function θ_n such that θ_n is bounded by $\sup_x|\varphi_n(x)|\leq n$, and θ_n is supported on $[-n-1,n+1]$ and $\{\theta_n\neq\varphi_n\}$ has arbitrarily small measure, in particular choose $m(\{\theta_n\neq\varphi_n\})\leq 2^{-n-1}$. Recall that given θ_n as a step function, there exists a continuous function $\widetilde{\theta_n}$ on $\mathbb{R}$ such that $\widetilde{\theta_n}$ is supported on $[-n-2,n+2]$ and $m(\{\widetilde{\theta_n}\neq\theta_n\})\leq 2^{-n-1}$. Thus $\{\ \widetilde{\{theta_n\}}$ is continuous and compactly supported, and

$$m(\{\widetilde{\theta_n}\neq\varphi_n\})\leq m(\{\widetilde{\theta_n}\neq\theta_n\})+m(\{\widetilde{\theta_n}\neq\varphi_n\})\leq 2^{-n}$$

and

$$\begin{aligned}
\|f-\widetilde{\theta_n}\|_p \leq \underbrace{\|f-\varphi_n\|_p}_{=0,\text{ as } n\to\infty} + \|\varphi_n-\widetilde{\theta_n}\|_p & \\
&\leq \left(\int_{\mathbb{R}} |\varphi_n-\widetilde{\theta_n}|^p\right)^{\frac{1}{p}} \\
&= \left(\int_{\widetilde{\theta_n}\neq\varphi_n} |\varphi_n-\widetilde{\theta_n}|^p\right)^{\frac{1}{p}} \\
&\leq ((2n)^p\cdot 2^{-n})^{\frac{1}{p}} \to 0, \text{as } n\to\infty
\end{aligned}$$

■

Here are some remarks:

① The density of $C_C(\mathbb{R})$ in $\mathscr{L}^p(\mathbb{R}$ is useful to study of properties of generic $\mathscr{L}^p$ functions, for example, if $f\in\mathscr{L}^p(\mathbb{R})$, then

$$\lim_{n\to\infty}\int_{\mathbb{R}}\left|f\left(x+\frac{1}{n}\right)-f(x)\right|^p dx=0$$

② However, $C_C(\mathbb{R})$ is not dense in $\mathscr{L}^\infty(\mathbb{R})$.

Convergence and Uniform Integrability

Recall the convergence theorems we studied earlier, suppose $\{f_n\}, f$ are all measurable and finite valued a.e functions, then:
① $f_n \to f$ in measure $\implies$ There exists a subsequence $\{n_k\}$ such that $f_{n_k} \to f$ a.e

② $f_n \to f$ a.e on A with $m(A) < +\infty \implies f_n \to f$ in measure on A.

So what about $f_n \to f$ in $\mathscr{L}^p(\mathbb{R}), 1 \leq p < +\infty$?

Theorem

Theorem 33. *If $\{f_n\}, f \subseteq \mathscr{L}^p(\mathbb{R}), 1 \leq p < +\infty$, and $f_n \to f$ in $\mathscr{L}^p(\mathbb{R})$, then $f_n \to f$ in measure.*

Proof. Choose $\forall \delta > 0$, $m(\{|f_n - f| > \delta\}) = \int_{\{|f_n - f| > \delta\}} 1dx$. Thus by monotoncity we have

$$\int_{\{|f_n-f|>\delta\}} 1dx \leq \int_{\{|f_n-f|>\delta\}} \frac{|f_n - f|^p}{\delta^p} dx$$

$$\leq \frac{1}{\delta^p} \int_{\mathbb{R}} |f_n - f|^p dx = \frac{1}{\delta^p} \|f_n - f\|_p^p \to 0, \text{as } n \to \infty.$$

∎

Remarks:
① In general, convergence in $\mathscr{L}^p(\mathbb{R})$ does not imply that convergence a.e

② When do we have that convergence a.e implies convergence in $\mathscr{L}^p(\mathbb{R})$? The answer is no in general, unless one of the integral convergence theorem applies.

③ When do we have convergence in measure applies convergence in $\mathscr{L}^p(\mathbb{R})$? The answer is still no in general, unless one of the integral convergence theorem applies.

Theorem

Theorem 34. *(mMON)*
Let $\{f_n\}, f$ be measurable and non-negative functions, and $\{f_n\}$ is increasing, also $f_n \to f$ in measure, then

$$\int_{\mathbb{R}} f = \lim_{n\to\infty} \int_{\mathbb{R}} f_n$$

Proof. $f_n \to f$ in measure implies $f_{n_k} \to f$ a.e along a chosen subsequence $\{n_k\}$. Now assume $mMON$ fails, then $\exists$ a subsequence $\{n_l\}$ such that

$$\int_{\mathbb{R}} f_{n_l} \not\to \int_{\mathbb{R}} f$$

However along f_{n_l} we have convergence in measure, then there exists a sub-sub sequence $\{n_{l_p}\}$ such that $f_{n_{l_p}} \to f$ a.e, thus

$$\lim_{p\to\infty} \int_{\mathbb{R}} f_{n_{l_p}} = \int_{\mathbb{R}} f$$

which is a contradiction. ■

Theorem

Theorem 35. *Let $\{f_n\}, f$ be measurable, $f_n \in \mathscr{L}^1(\mathbb{R})$ and $f_n \to f$ in measure. If $\exists g \in \mathscr{L}^1(\mathbb{R})$ such that $|f_n| \le |g|$ for all $n \ge 1$, then $f_n \to f$ in $\mathscr{L}^1(\mathbb{R})$.*

Proof. :) ■

Proposition

Proposition 61. *Given $A \in \mathscr{M}$ with $m(A) < +\infty$, then $f \in \mathscr{L}^1(A)$ if and only if*

$$\lim_{n\to\infty} \int_{A\cap\{|f|>n\}} |f| = 0$$

Proof. $(\Longrightarrow)$: Trivial;

$(\Longleftarrow)$: Choose a N large enough such that $\displaystyle\int_{A\cap\{|f|>N\}} |f| \le 1$, then

$$\begin{aligned}\int_A |f| &= \int_{A\cap\{|f|\le N\}} |f| + \int_{A\cap\{|f|>N\}} |f| \\ &\le M\cdot m(A) + 1 < +\infty.\end{aligned}$$

■

Definition

Definition 25. *Given a sequence $\{f_n\}$ measurable and $A \in \mathscr{M}$, then say that $\{f_n\}$ is uniformly integrable on A if*

$$\lim_{M\to\infty} \sup_{n\ge 1} \int_{A\cap\{|f_n|>M\}} |f_n| = 0$$

Proposition

Proposition 62. *Let $\{f_n\}$ be measurable and $A \in \mathscr{M}$, then*

① *If $m(A) < +\infty$, and $\{f_n\}$ is uniformly integrable on A, then $\{f_n\}$ is bounded on $\mathscr{L}^1(A)$, i.e*

$$\sup_{n\ge 1} \int_A |f_n| < +\infty.$$

② *If $\{f_n\}$ is bounded in $\mathscr{L}^p(A)$ for all $1 < p < +\infty$, then $\{f_n\}$ is uniformly integrable on A.*

Proof. ① :
Choose M to be large enough such that

$$\sup_{n\geq 1}\int_{A\cap\{|f_n|>M\}}|f_n|<1$$

thus

$$\sup_{n\geq 1}\int_A|f_n|=\sup_{n\geq 1}\int_{A\cap\{|f_n|\leq M\}}|f_n|+\sup_{n\geq 1}\int_{A\cap\{|f_n|>M\}}|f_n|$$
$$\leq M\cdot m(A)+1<+\infty.$$

■

Proof. ②:
For any $M>0$, we have

$$\int_{A\cap\{|f_n|>M\}}|f_n|\leq\int_{A\cap\{|f_n|>M\}}|f_n|\cdot\frac{|f_n|^{p-1}}{M^{p-1}}$$

$$\leq\frac{1}{M^{p-1}}\int_A|f_n|^p.$$

Since $\{f_n\}$ is bounded in $\mathscr{L}^p(A)$, then

$$\lim_{M\to\infty}\frac{1}{M^p}\sup_{n\geq 1}\int_A|f_n|^p=0.$$

■

Proposition

Proposition 63. *Given $\{f_n\}$ measurable, $A\in\mathscr{M}$ and $m(A)<+\infty$, TFAE:*

① $f_n\in\mathscr{L}^1(A),\forall n\geq 1$ *and* $f_n\to f$ *in* $\mathscr{L}^1(A)$.

② $\{f_n\}$ *is uniformly integrable on A and $f_n\to f$ in measure on A.*

Proof. ① implies ② :
Assume $f_n\to f$ in $\mathscr{L}^1(A)$, thus $\int_A|f_n|\to\int_A|f|$, so $\{f_n\}$ is bounded in $\mathscr{L}^1(A)$ and $\forall M>0$, then we have

$$\int_{A\cap\{|f_n|>M\}}|f_n|\leq\underbrace{\int_{A\cap\{|f_n|>M\}}|f_n-f|}_{\leq\int_A|f_n-f|\to 0,}+\int_{A\cap\{|f_n|>M\}}|f|,$$

and furthermore

$$\int_{A\cap\{|f_n|>M\}}|f|=\int_{A\cap\{|f_n|>M\}\cap\{|f|\leq\sqrt{M}\}}|f|+\int_{A\cap\{|f_n|>M\}\cap\{|f|>\sqrt{M}\}}|f|$$
$$\leq\sqrt{M}\cdot m(A\cap\{|f_n|>M\})+\int_{A\cap\{|f|>\sqrt{M}\}}|f|$$
$$\leq\left(\sqrt{M}\frac{\sup_n\int_A|f_n|}{M}+\int_{A\cap\{|f|>\sqrt{M}\}}|f|\right)\longrightarrow 0,\text{as }M\to\infty.$$

So we see that $\forall \varepsilon > 0$, choose N large enough, such that $\forall n \geq N$, we have

$$\int_A |f_n - f| \leq \frac{\varepsilon}{3},$$

then choose M large enough as well, such that $\displaystyle\int_{A\cap\{|f|>\sqrt{M}\}} |f| < \frac{\varepsilon}{3}$ and $\dfrac{\sup_n \int_A |f_n|}{\sqrt{M}} < \dfrac{\varepsilon}{3}$, so

$$\sup_{n\geq N} \int_{A\cap\{|f_n|>M\}} |f_n| \leq \varepsilon.$$

By making M even larger if necessary, such that we have

$$\int_{A\cap\{|f_k|>M\}} |f_n| < \varepsilon, \forall k = 1, 2, \cdots, N-1.$$

② implies ①:
:)

∎

Chapter 3

Product Spaces

Product σ-Algebra

In this whole chapter, we will only demonstrate the construction of product space with $\mathbb{R}^2$, but the same construction can be generalized to $\mathbb{R}^d$.

Definition

Definition 26. *Consider $\mathbb{R}^2$, a product space, $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ with elements (x,y), the product σ-algebra of subsets of $\mathbb{R}^2$, denoted by $\mathscr{M}^2$, is defined as*

$$\mathscr{M}^2 := \sigma(\{A \times B : A, B \in \mathscr{M}\})$$

We call $A \times B$ is a product set

$$A \times B := \{(x,y) : x \in A, y \in B\}$$

$\mathscr{M}^2$ contains all the rectangles $I_1 \times I_2$ where I_1, I_2 are intervals, also it contains all the singletons, all the open sets and closed sets, or we can say

$$\mathfrak{B}(\mathbb{R}^2) := \sigma(\{\text{open sets in } \mathbb{R}^2\})$$

Definition

Definition 27. *Given $E \subseteq \mathbb{R}^2$, we define:*

① $\forall x \in \mathbb{R}$, $E_x := \{y \in \mathbb{R} : (x,y) \in E\}$*, is called the slice of E at x;*
② $\forall y \in \mathbb{R}$, $E^y := \{x \in \mathbb{R} : (x,y) \in E\}$*, is called the slice of E at y.*

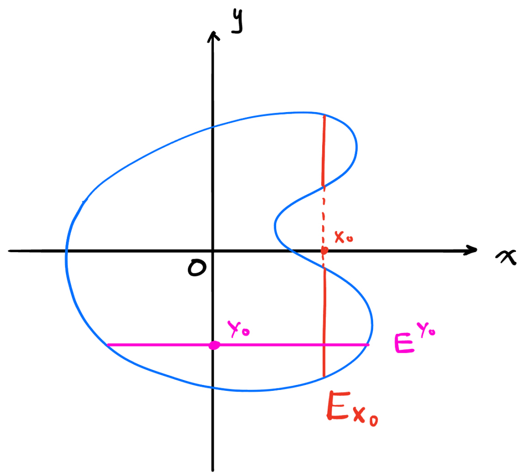


Figure 3.1: A geometric view of E_x and E^y

Proposition

Proposition 64. *If $E \in \mathscr{M}^2$, then $\forall x \in \mathbb{R}, E_x \in \mathscr{M}, \forall y \in \mathbb{R}, E^y \in \mathscr{M}$.*

In words, product measurability implies marginal measurability.

Proof. Define $\mathscr{A} = \{E \subseteq \mathbb{R}^2 : E_x \in \mathscr{M}, \forall x \in \mathbb{R}\}$, then we have:
① $\mathbb{R}^2 \in \mathscr{A}$, since $\mathbb{R}^2_x = \mathbb{R} \in \mathscr{M}, \forall x$;

② If $E \in A$, then $\forall x \in \mathbb{R}, (E^C)_x := \{y : (x,y) \in E\} = (E_x)^C \in \mathscr{M}$;
③ Suppose a sequence $\{E_n\} \subseteq \mathscr{A}$, then $\forall x \in \mathbb{R}$, $\left(\bigcup_{n=1}^{+\infty} E\right)_x = \{y : (a,y) \in E_n \text{ for some } n\ \} = \bigcup_{n=1}^{+\infty}(E_n)_x \in \mathscr{M}$.

So we claim that $\mathscr{A}$ is a σ-algebra of subsets of $\mathbb{R}$. So $\forall A, B \in \mathscr{M}$, we have

$$(A \times B)_x := \begin{cases} B : x \in A \\ \varnothing : \text{otherwise} \end{cases} \in \mathscr{M} \implies A \times B \in \mathscr{A} \implies \mathscr{M}^2 \subseteq \mathscr{A}$$

Therefore, $\forall E \in \mathscr{M}^2$, we have $E_x \in \mathscr{M}$ for all x. Same statement can be drawn on E^y using the method above.

■

In fact, there exists $E \subseteq \mathbb{R}^2$ such that $\forall x \in \mathbb{R}$, $E_x \in \mathscr{M}$ and $\forall y \in \mathscr{M}, E^y \in \mathscr{M}$ but $E \notin \mathscr{M}^2$. This construction is done by Sierpiński.

Definition

Definition 28. *Assume $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ is a function, then $\forall x \in \mathbb{R}$, $f_x : \mathbb{R} \to \overline{\mathbb{R}}$ s.t $f_x(y) = f(x,y), \forall y$ (the slice of f at x, and likewise $f^y : \mathbb{R} \to \overline{\mathbb{R}} : f^y(x) = f(x,y) \forall x$ (the slice of f at y).*

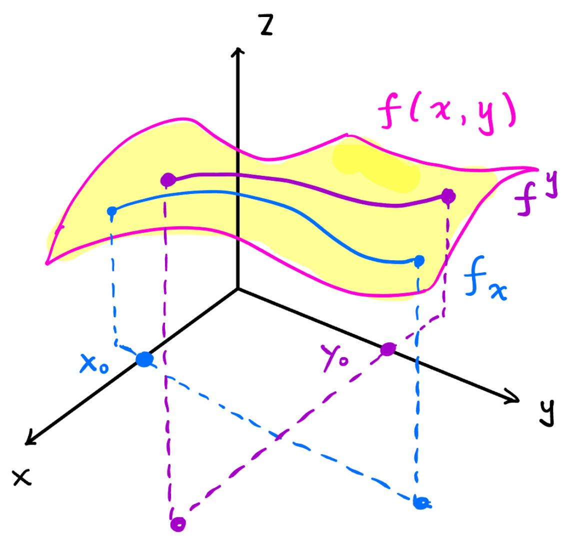


Figure 3.2: A geometric view of f_x and f^y

Definition

Definition 29. *A function* $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ *is* $\mathscr{M}^2$*-measurable, if* $\forall a \in \mathbb{R}, f^{-1}([-\infty, a)) \in \mathscr{M}^2$.

Proposition

Proposition 65. $f : \mathbb{R}^2 \to \mathbb{R}^2$ *is* $\mathscr{M}^2$ *measurable, then* $\forall x \in \mathbb{R}$, $f_x : \mathbb{R} \to \overline{\mathbb{R}}$ *is* $\mathscr{M}$ *measurable and* $\forall y \in \mathbb{R}$ *:* $f^y : \mathbb{R} \to \overline{\mathbb{R}}$ *is* $\mathscr{M}$ *measurable.*

Proof. We first note that $\forall B \subseteq \mathbb{R}^2$, we have

$$(f^{-1}(B))_x = f_x^{-1}(B), \forall x \in \mathbb{R} \quad \text{and} \quad (f^{-1}(B))^y = (f^y)^{-1}(B), \forall y \in \mathbb{R}$$

Thus $\forall a \in \mathbb{R}$, $f_x^{-1}([-\infty, a)) = (f^{-1}([-\infty, a))_x \in \mathscr{M}$, and same for y.

■

Remarks:

① : If $f : \mathbb{R}^2 \to \mathbb{R}$ is continuous, then f is $\mathscr{M}^2$ measurable;

② : If $f = \mathbf{1}_E$ for some $E \subseteq \mathbb{R}^2$, then f is $\mathscr{M}^2$ measurable if and only if $E \in \mathscr{M}^2$;

③ f_x is $\mathscr{M}$-measurable for all $x \in \mathbb{R}$ **DOES NOT** generally imply f is $\mathscr{M}^2$ measurable;

④ If $f(x, y) = h(x)g(y)$ for some non-trivial functions $h, g : \mathbb{R} \to \overline{\mathbb{R}}$, then f is $\mathscr{M}^2$ measurable if and only if h, g are both $\mathscr{M}$-measurable.

Product Measure

Definition

Definition 30. *Given $E \in \mathscr{M}^2$, define functions $I_E^1 : \mathbb{R} \to \overline{\mathbb{R}}$ by $\forall x \in \mathbb{R}, I_E^1(x) := m(E_x)$ and $I_E^2 : \mathbb{R} \to \overline{\mathbb{R}}$ by $\forall y \in \mathbb{R}, I_E^2(y) := m(E^y)$.*

Theorem

Theorem 36. *Given $E \in \mathscr{M}^2$ and claim that I_E^1, I_E^2 are $\mathscr{M}$ measurable functions, then*

$$\int_{\mathbb{R}} I_E^1(x)dx = \int_{\mathbb{R}} I_E^2(y)dy \quad (*)$$

Are you ready for the longest proof in this course so far?

Proof. If indeed I_E^1, I_E^2 are $\mathscr{M}$ measurable, then $(*)$ is well defined, since I_E^1, I_E^2 are non-negative, set

$$\Sigma := \{E \in \mathscr{M}^2, \text{all the above hold for } E_N = E \cap [-N,N]^2 \ \}$$

i.e

$$\Sigma := \{E \in \mathscr{M}^2, I_{E_N}^1, I_{E_N}^2 \ \text{ are measurable and (*) holds}\}$$

Note that $\forall E \in \mathscr{M}^2, N > 0, I_{E_N}^1(x) = \begin{cases} m((E_N)_x), x \in [-N,N] \\ 0, \text{otherwise} \end{cases}$ and likewise we also have $I_{E_N}^y = \begin{cases} m((E_N)^y), y \in [-N,N] \\ 0, \text{otherwise} \end{cases}$, and furthermore

$$I_{E_N}^1(x) = \mathbf{1}_{[-N,N]} I_{E_N}(x)$$

$$I_{E_N}^2(x) = \mathbf{1}_{[-N,N]} I_{E_N}^2(y)$$

We denote

$$\mathscr{C} := \{A \times B, A, B \in \mathscr{M}\}, \mathscr{M}^2 = \sigma(\mathscr{C})$$

CLAIM 1 : $\mathscr{C} \subseteq \Sigma$.

$\forall E \in \mathscr{C}$, $\exists A, B \in \mathscr{M}$ s.t $E = A \times B$, and $\forall N > 0$, $E_N = (A \times B) \cap [-N,N]^2$, we also denote $A_N = A \cap [-N,N]; B_N = B \cap [-N,N]$, now we have

$$I_{E_N}^1(x) = I_{A_N \times B_N}^1(x) = \begin{cases} m(B_N), x \in A_N \\ 0, \text{otherwise} \end{cases}$$

and

$$I_{E_N}^2(y) = I_{A_N \times B_N}^2(y) = \begin{cases} m(A_N), y \in B_N \\ 0, \text{otherwise} \end{cases}$$

So it is clear that $I^1_{A_N\times B_N}, I^2_{A_N\times B_N}$ are $\mathscr{M}$ measurable and

$$\int_{\mathbb{R}} I^1_{A_N\times B_N}(x)dx = m(B_N)\cdot m(A_N) = \int_{\mathbb{R}} I^2_{A_N\times B_N}(y)dy$$

which means $(*)$ holds, and thus $\mathscr{C}\subseteq\Sigma$.

CLAIM 2 : $\mathbb{R}^2\in\Sigma$.

$\forall N>0$, we know that

$$I^1_{[-N,N]^2}(x) = \begin{cases} 2N, x\in[-N,N] \\ 0, \text{otherwise} \end{cases}$$

and

$$I^2_{[-N,N]^2}(y) = \begin{cases} 2N, y\in[-N,N] \\ 0, \text{otherwise} \end{cases}$$

Then clearly $I^1_{[-N,N]^2}(x), I^2_{[-N,N]^2}(y)$ are $\mathscr{M}$ measurable and $(*)$ holds trivially, which is $4N^2$.

CLAIM 3: If $E\in\Sigma$, then $E^C\in\Sigma$.

$\forall N>0$, denote by $F_N = E^C\cap[-N,N]^2$. Obviously $I^1_{F_N}(x)=0$ if $x\notin[-N,N]$ and if $x\in[-N,N]$, $(F_N)_x = \{y:(x,y)\in E^C\cap[-N,N]^2\} = [-N,N]\backslash E_x = [-N,N]\backslash(E_N)_x$, which means

$$I^1_{F_N}(x) = 2N - I^1_{E_N}(x)$$

Likewise we can do the same for $I^2_{F_N}(y)$, that is,

$$I^2_{F_N}(y) = \begin{cases} 2N - I^2_{E_N}(y), y\in[-N,N] \\ 0, \text{otherwise} \end{cases}$$

So again, $I^1_{F_N}(x); I^2_{F_N}(y)$ are measurable, as for their integrals, we have

$$\int_{\mathbb{R}} I^1_{F_N}(x)dx = \int_{[-N,N]}(2N - I^1_{E_N}(x))dx = 4N^2 - \int_{\mathbb{R}} I^1_{E_N}(x)dx$$

and

$$\int_{\mathbb{R}} I^2_{F_N}(y)dy = \int_{[-N,N]}(2N - I^2_{E_N}(y))dy = 4N^2 - \int_{\mathbb{R}} I^2_{E_N}(y)dy$$

thus $(*)$ holds trivially.

CLAIM 4: Σ is closed under taking countable union.

W.L.O.G, assume that $\{E_K\}\subseteq\Sigma$ are disjoint, now for each fixed $N>0$, let

$$E_N = \bigcup_{k=1}^{+\infty} E_{kN}; E_{kN} = E_K\cap[-N,N]^2$$

therefore

$$\begin{aligned} I^1_{E_N}(x) &= \mathbf{1}_{[-N,N]}(x) \cdot m\left(\bigcup_{k=1}^{+\infty}(E_{kN})_x\right) \\ &= \mathbf{1}_{-N,N]}(x) \cdot \sum_{k=1}^{+\infty} m((E_{kN})_x) \\ &= \sum_{k=1}^{+\infty} I^1_{E_{kN}}(x) \end{aligned}$$

So likewise,

$$I^2_{E_N}(y) = \sum_{k=1}^{+\infty} I^2_{E_{kN}}(y)$$

and we have $I^1_{E_N}, I^2_{E_N}$ both measurable, moreover

$$\int_{\mathbb{R}} I^1_{E_N}(x)dx = \sum_{k=1}^{+\infty} \int_{\mathbb{R}} I^1_{E_{kN}}(x)dx$$

and

$$\int_{\mathbb{R}} I^2_{E_N}(y)dy = \sum_{k=1}^{+\infty} \int_{\mathbb{R}} I^2_{E_{kN}}(y)dy$$

which means $(*)$ holds trivially, thus Σ i closed under taking countable union.

CLAIM 5: Σ is a σ-algebra of subsets of $\mathbb{R}^2$, also $\mathscr{C} \subseteq \Sigma$, so $\Sigma = \mathscr{M}^2$.

This claim is easy to check based on our construction of the previous claims.

Now, $\forall E \in \mathscr{M}^2$, $E \in \Sigma$. So all the statements are true for E_N, we have

$$\begin{aligned} I^1_E(x) &= \lim_{N\to\infty} \mathbf{1}_{[-N,N]}(x) \cdot m((E_N)_x) \\ &= \lim_{N\to\infty} I^1_{E_N}(x) \end{aligned}$$

Moreover $I^1_{E_N}(x)$ is increasing, thus $I^1_E(x)$ is measurable, so

$$\int_{\mathbb{R}} I^1_E(x)dx = \lim_{N\to\infty} \int_{\mathbb{R}} I^1_{E_N}(x)dx$$

Likewise, I^2_E is measurable, and

$$\int_{\mathbb{R}} I^2_E(y)dy = \lim_{N\to\infty} \int_{\mathbb{R}} I^2_{E_N}(y)dy$$

and $(*)$ holds trivially.

$\blacksquare$

Definition

Definition 31. *We define a non-negative set function on the measurable space* $(\mathbb{R}^2, \mathscr{M}^2)$ *by:*

$$\forall E \in \mathscr{M}^2, m(E) := \int_{\mathbb{R}} I_E^1(x)dx = \int_{\mathbb{R}} I_E^2(y)dy$$

this function m is called the Lebesgue measure on $\mathbb{R}^2$.

Indeed, m is a measure on $(\mathbb{R}^2, \mathscr{M}^2)$, i.e

① $m(\varnothing) = 0$;

② If $\{E_k\} \subseteq \mathscr{M}^2$ are disjoint, then

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

Here are some remarks:

① For any rectangle $E = I_1 \times I_2$, $m(E) = \mathbf{Area}(I_1 \times I_2)$, thus any singleton, countable set, line segment are null set in $\mathbb{R}^2$;

② If $A \subseteq \mathbb{R}$ is a null set under the Lebesgue measure on $\mathbb{R}$, then $A \times \mathbb{R}, \mathbb{R} \times A$ are again null sets under the Lebesgue measure on $\mathbb{R}^2$;

③ $\mathscr{M}^2$ is not complete under the Lebesgue measure on $\mathbb{R}^2$, e.g suppose N is the Vitalli set on $\mathbb{R}$ (non-measurable), and $a \in \mathbb{R}$, by definition the set $\{a\} \times N$ should be of the measure 0, which leads a contradiction;

④ It is possible to construct Lebesgue measure on $\mathbb{R}^2$ through the approach using outer measure, we denote $\overline{\mathscr{M}^2}$ to be the completion of $\mathscr{M}^2$ under the Lebesgue measure;

⑤ The Lebesgue measure m on $\mathbb{R}^2$ is the **ONLY** measure on $\mathscr{M}^2$, $\mathfrak{B}(\mathbb{R}^2), \overline{\mathscr{M}^2}$ such that for all rectangle $E = I_1 \times I_2$, $m(E) = m(I_1) \cdot m(I_2)$. Because in fact $\mathscr{I} := \{I_1 \times I_2, I_1, I_2 \text{ are finite}\}$ is a π-system and $\sigma(\mathscr{I}) = \mathfrak{B}(\mathbb{R}^2)$;

⑥ The Lebesgue measure m on $\mathbb{R}^2$ is translation invariant;

⑦ The Lebesgue measure m on $\mathbb{R}^2$ is the **only** measure on $\mathscr{M}^2$, $\mathfrak{B}(\mathbb{R}^2), \overline{\mathscr{M}^2}$ such that it's translation invariant and assigns 1 to the unit rectangle $[0,1]^2$.

Fubini-Tonelli Theorem

Definition

Definition 32. *Let $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ be $\mathscr{M}^2$ measurable and non-negative, then we define*

$$I_f^1(x) := \int_{\mathbb{R}} f(x,y)dy = \int_{\mathbb{R}} f_x(y)dy, \forall x \in \mathbb{R}$$

$$I_f^2(y) := \int_{\mathbb{R}} f(x,y)dx = \int_{\mathbb{R}} f^y(x)dx, \forall y \in \mathbb{R}$$

Given $f : \mathbb{R}^2 \to [0,+\infty]$ to be $\mathscr{M}^2$ measurable and non-negative, the integral of f with respect to m, the Lebesgue measure on $\mathbb{R}^2$ is denoted by

$$\int_{\mathbb{R}^2} f(x,y)dxdy = \int_{\mathbb{R}^2} f$$

Theorem

Theorem 37. *(Tonelli's Theorem)*

Let $f : \mathbb{R}^2 \to [0,+\infty]$ be $\mathscr{M}^2$ measurable and non-negative, then

$$\int_{\mathbb{R}^2} f(x,y) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)dy \right) dx = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)dx \right) dy$$

Proof. Since f is $\mathscr{M}^2$ measurable and non-negative, then there exists a sequence of simple functions $\{\varphi_n\}$ such that $\{\varphi_n\}$ is increasing and $\lim_{n\to\infty} \varphi_n = f$ and

$$\int_{\mathbb{R}^2} f = \lim_{n\to\infty} \int_{\mathbb{R}^2} \varphi_n$$

where we denote

$$\varphi_n = \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbf{1}_{A_{n,k}} \ \ A_{n,k} := \left\{ (x,y) \in [-n,n]^2 : \frac{k}{2^n} \leq f(x,y) < \frac{k+1}{2^n} \right\} \in \mathscr{M}^2$$

and thus

$$\int_{\mathbb{R}^2} f(x,y)dxdy = \lim_{n\to\infty} \sum_{k=0}^{n2^n} \frac{k}{2^n} m(A_{n,k}) \qquad (*)$$

Now $\forall x \in \mathbb{R}$,

$$\begin{aligned}
I_f^1(x) = \int_{\mathbb{R}} f(x,y)dy &= \lim_{n\to\infty} \int_{\mathbb{R}} \varphi_n(x,y)dy \\
&= \lim_{n\to\infty} \sum_{k=0}^{n2^n} \frac{k}{2^n} \underbrace{I_{A_{n,k}}^1(x)}_{m((A_{n,k})_x)} \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} \underbrace{\sum_{k=0}^{n2^n} \frac{k}{2^n} I_{A_{n,k}}^1(x)dx}_{I_{\varphi_n}^1(x)} \\
&= \lim_{n\to\infty} \sum_{k=0} \frac{k}{2^n} \int_{\mathbb{R}} I_{A_{n,k}}^1(x)dx \qquad (**)
\end{aligned}$$

Now likewise,

$$\int_{\mathbb{R}} I_f^2(y)dy = \lim_{n\to\infty} \sum_{k=0}^{n2^n} \int_{\mathbb{R}} I_{A_{n,k}}^2(y)dy \qquad (***)$$

Now compare $(*), (**), (***)$, and by definition of m on $\mathbb{R}^2$,

$$m(A_{n,k}) = \int_{\mathbb{R}} I_{A_{n,k}}^1(x)dx = \int_{\mathbb{R}} I_{A_{n,k}}^2(y)dy$$

and thus it implies

$$\int_{\mathbb{R}^2} f(x,y)dxdy = \int_{\mathbb{R}} I_f^1(x)dx = \int_{\mathbb{R}} I_f^2(y)dy$$

■

Definition

Definition 33. *Given $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ is $\mathscr{M}^2$ measurable, we say $f \in \mathscr{L}^1(\mathbb{R}^2)$, if*

$$\int_{\mathbb{R}^2} |f(x,y)|dxdy < +\infty$$

Also we know that

$$\int_{\mathbb{R}^2} f(x,y)dxdy = \int_{\mathbb{R}^2} f^+(x,y)dxdy - \int_{\mathbb{R}^2} f^-(x,y)dxdy$$

Assume $f \in \mathscr{L}^1(\mathbb{R}^2$, thus by Tonelli's Theorem, we have

$$\int_{\mathbb{R}^2} |f(x,y)|dxdy = \int_{\mathbb{R}} \left[\underbrace{\int_{\mathbb{R}} |f(x,y)|dx}_{I_{|f|}^1 \in \mathscr{L}^1(\mathbb{R}} \right] dy = \int_{\mathbb{R}} \left[\underbrace{\int_{\mathbb{R}} |f(x,y)|dy}_{I_{|f|}^2 \in \mathscr{L}^1(\mathbb{R}} \right] dx$$

Theorem

Theorem 38. *(Fubini's Theorem)*

If $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ is $\mathscr{M}^2$ measurable and f is integrable, then:

① $I_f^1, I_f^2 \in \mathscr{L}^1(\mathbb{R})$;

② $I_f^1(x)$ *is finite-valued for a.e x, and* $f_x \in \mathscr{L}^1(\mathbb{R})$ *for a.e x ;* $I_f^2(y)$ *is finite valued for a.e y and* $f_y \in \mathscr{L}^1(\mathbb{R})$ *for a.e y;*

③ *We have*

$$\int_{\mathbb{R}^2} dxdy = \int_{\mathbb{R}} I_f^1(x)dx = \int_{\mathbb{R}} I_f^2(y)dy$$

Proof. Assume that $f \in \mathscr{L}^1(\mathbb{R}^2)$, then

$$\begin{aligned}\int_{\mathscr{R}^2} |I_f^1(x)|dx &\le \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x,y)|dy \right) dx \\ &= \int_{\mathbb{R}} |f(x,y)|dxdy < +\infty\end{aligned}$$

thus $I_f^1 \in \mathscr{L}^1(\mathbb{R})$ and likewise, $I_f^2 \in \mathscr{L}^1(\mathbb{R})$. Thus (ii) is trivial given the fact on (i). As for (iii), write $f = f^+ - f^-$, and $f \in \mathscr{L}^1(\mathbb{R}) \implies f^+, f^- \in \mathscr{L}^1(\mathbb{R})$, thus f^+, f^- are finite valued a.e, then according to Tonelli's Theorem,

$$\int_{\mathbb{R}} I_{f^+}^1(x)dx = \int_{\mathbb{R}} I_{f^+}^2(y)dy = \int_{\mathbb{R}^2} f^+(x,y)dxdy < +\infty \qquad (*)$$

$$\int_{\mathbb{R}} I_{f^-}^1(x)dx = \int_{\mathbb{R}} I_{f^-}^2(y)dy = \int_{\mathbb{R}^2} f^-(x,y)dxdy < +\infty \qquad (**)$$

Thus $I_{f^+}^1, I_{f^-}^1, I_{f^+}^2, I_{f^-}^2 \in \mathscr{L}^1(\mathbb{R})$ and hence are finite valued a.e, thus by linearity,

$$\int_{\mathbb{R}} I_{f^+}^1(x) - \int_{\mathbb{R}} I_{f^-}^1(x)dx = \int_{\mathbb{R}} \left(I_{f^+}^1(x) - I_{f^-}^1(x) \right) dx$$

for a.e $x \in \mathbb{R}$, $f_x^+ \in \mathscr{L}^1(\mathbb{R})$ and $f_x^-(x) \in \mathscr{L}^1(\mathbb{R})$, by linearity,

$$\begin{aligned}I_{f^+}^1(x) - I_{f^-}^1(x) &= \int_{\mathbb{R}} f_x^+(y)dy - \int_{\mathbb{R}} f_x^-(y)dy \\ &= \int_{\mathbb{R}} \left(\underbrace{f_x^+(y) - f_x^-(y)}_{f_x} \right) dy\end{aligned}$$

thus

$$\int_{\mathbb{R}} I_{f^+}^1(x)dx - \int_{\mathbb{R}} I_{f^-}^1(x)dx = \int_{\mathbb{R}} I_f^1(x)dx$$

Likewise we also have

$$\int_{\mathbb{R}} I_{f^+}^2(y)dy - \int_{\mathbb{R}} I_{f^-}^2(y)dy = \int_{\mathbb{R}} I_f^2(y)dy$$

Now taking $(*)-(**)$, we have

$$\int_{\mathbb{R}} I_f^1(x)dx - \int_{\mathbb{R}} I_f^2(y)dy = \int_{\mathbb{R}^2}(f^+ - f^-) = \int_{\mathbb{R}^2} f$$

■

Remarks:

① In general $I_f^1, I_f^2 \in \mathscr{L}^1(\mathbb{R})$ **DO NOT** imply $f \in \mathscr{L}^1(\mathbb{R})$. For example, construct

$$f(x,y) = \begin{cases} 1, x < y < x+1 \\ -1, x-1 < y < x \\ 0, \text{otherwise} \end{cases}$$

then $I_f^1(x) = I_f^2(y) = 0$, but $f \notin \mathscr{L}^1(\mathbb{R}^2)$;

② If $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ is $\overline{\mathscr{M}^2}$ measurable, then Fubini and Tonelli's Theorem still apply;

③ The construction can also be extended to $\mathbb{R}^d$, $d \geq 3$.

Chapter 4

Differentiation Theory

In the Riemann setting, the notation of differentiation is closely related to integration, for example, if

$$F(x) = \int_a^x f(t)dt$$

for $x \in [a,b]$ and f being Riemann integrable on $[a,b]$, then F is differentiable and we have

$$F'(x) = f(x), x \in (a,b)$$

Also, if F is differentiable and F' is Riemann integrable on $[a,b]$, then

$$F(b) - F(a) = \int_a^b F'(x)dx$$

What if we consider the Lebesgue setting?

Hardy-Littlewoood (HL) maximal function

Definition

Definition 34. *Assume $f \in \mathscr{L}^1(\mathbb{R})$, the Hardy-Littlewood (HL) maximal function of f, denoted by f^*, is defined as*

$$\forall x \in \mathbb{R}, f^*(x) := \sup_{I \in \mathscr{I}(x)} \left\{ \frac{1}{m(I)} \int_I |f(t)|dt \right\}$$

where

$$\mathscr{I}(x) := \{\text{open interval } I \text{, such that } x \in I.\}$$

Proposition

Proposition 66. *Given $f \in \mathscr{L}^1(\mathbb{R})$, let f^* be the HL maximal function of f, then f^* is $\mathscr{M}$ measurable.*

Proof. It is sufficient to show that $\forall a \in [0,+\infty]$, $\{f^* > a\} \in \mathscr{M}$. If $x \in \{f^* > a\}$, thus $f^*(x) > a$ and by definition,

$$a \leq f^*(x) = \sup_{I \in \mathscr{I}(x)} \left\{ \frac{1}{m(I)} \int_I |f(x)|dx \right\}$$

which implies $\exists I \in \mathscr{I}(x)$, such that $\frac{1}{m(I)} \int_I |f(x)|dx > a$, and since I is open, so $\exists \delta > 0$, such that $(x-\delta, x+\delta) \subseteq I$, i.r $\forall y \in (x-\delta, x+\delta), y \in I$, $I \in \mathscr{I}(y)$, which means

$$\begin{aligned} f^*(y) &= \sup_{J \in \mathscr{I}(y)} \left\{ \frac{1}{m(J)} \int_J |f(y)|dy \right\} \\ &\geq \frac{1}{m(I)} \int_I |f(x)|dx > a \end{aligned}$$

which means that $y \in \{f^* > a\}$, thus

$$(x-\delta, x+\delta) \subseteq \{f^* > a\}$$

This implies that $\{f^* > a\}$ is open, thus measurable. ■

Recall that if $f \in \mathscr{L}^1([a,b])$, then we define $F(x) = \int_{[a,x]} f(t)dt$, since a singleton has measure 0, we shall write $F(x)$ as

$$F(x) = \int_a^x f(t)dt, \forall x \in [a,b]$$

If we are interested in $F'(x)$, which we will define later as the derivative of $F(x)$, is of the form

$$\lim_{h \to 0} \frac{F(x+h) - F(x)}{h}$$

also

$$F'(x) = \lim_{h \to 0} \frac{1}{h} \int_x^{x+h} f(t)dt = \lim_{h \to 0} \frac{1}{m((x,x+h))} \int_{[x,x+h]} f(t)dt$$

Theorem

Theorem 39. *(Vitali's Covering Lemma)*

Assume that $\mathscr{I} := \{I_1, I_2, \cdots, I_N\}$ is a finite sequence of open intervals, then there exists a sub-collection of intervals

$$J := \{I_{k_1}, I_{k_2}, \cdots, I_{k_M}\} \subseteq \mathscr{I}$$

such that $I_{k_i} \bigcap I_{k_j} = \varnothing$ for $i \neq j$ and

$$m\left(\bigcup_{n=1}^N I_n\right) \leq 3 \sum_{j=1}^M m(I_{k_j})$$

Proof. $W.L.O.G$, assume that $m(I_n) < +\infty$ for every $I_n \in \mathscr{I}$, we begin with the biggest interval in $\mathscr{I}$, denote that interval by I_{k_1}, and set

$$\mathscr{I}_{k_1} := \{I \in \mathscr{I}, \text{such that } I \bigcap I_{k_1} \neq \varnothing\}$$

We claim that $\forall I \in \mathscr{I}_{k_1}$, $I \cap I_{k_1} \neq \varnothing$ and $m(I) \leq m(I_{k_1})$ and $I \subseteq 3I_{k_1}$, and this implies that

$$\bigcup_{I \in \mathscr{I}_{k_1}} I \subseteq 3I_{k_1}$$

Now consider $\mathscr{I} \backslash \mathscr{I}_{k_1}$ and we denote the biggest interval among those to be I_{k_2}, and set

$$\mathscr{I}_{k_2} := \{I \in \mathscr{I} \backslash \mathscr{I}_{k_1}, \text{such that } I \bigcap I_{k_2} \neq \varnothing\}$$

and likewise we have

$$\bigcup_{I \in \mathscr{I}_{k_2}} I \subseteq 3I_{k_2}$$

So by repeating this process until all intervals have used, and by choice we know $I_{k_i} \bigcap I_{k_j} = \varnothing$ if $k_i \neq k_j$, and $\mathscr{I}_{k_1}, \mathscr{I}_{k_2}, \cdots, \mathscr{I}_{k_M}$ forms a partition of $\mathscr{I}$, then

$$m\left(\bigcup_{n=1}^{N} I_n\right) = \sum_{j=1}^{M} m\left(\bigcup_{I \in \mathscr{I}_{k_j}} I\right) \leq 3 \sum_{j=1}^{M} (I_{k_j}).$$

■

Proposition

Proposition 67. *Suppose $f \in \mathscr{L}^1(\mathbb{R})$ and f^* be the HL maximal function of f, then $\forall \varepsilon > 0$, we have*

$$m(\{x \in \mathbb{R} : f^*(x) > \varepsilon\}) \leq \frac{3}{\varepsilon} \|f\|_1 := \frac{3}{\varepsilon} \int_{\mathbb{R}} |f|$$

Proof. Fix $\varepsilon > 0$, write $B := \{x \in \mathbb{R} : f^*(x) > \varepsilon\}$, recall that

$$m(B) := \sup\{m(K) : K \subseteq B, K \text{ is compact}\}$$

Thus it suffices to show that $m(K) \leq \frac{3}{\varepsilon}\|f\|_1$ for every compact set $K \subseteq B$. For every $x \in L \subseteq B$, we know that $f^*(x) > \varepsilon$, thus there exists open interval I_x such that $x \in I_x$ and also $\dfrac{1}{m(I_x)} \displaystyle\int_{I_x} |f| > \varepsilon$, then $K \subseteq \bigcup_{x \in K} I_x$, and there exists $\mathscr{I} := \{I_1, I_2, \cdots, I_N\}$ such that $K \subseteq \bigcup_{i=1}^{N} I_i$, and by the Canteli covering lemma, we now have

$$m(K) \leq m\left(\bigcup_{i=1}^{N} I_i\right) \leq 3 \sum_{j=1}^{M} m(I_{k_j})$$

for some disjoint sets $I_{k_1}, I_{k_2}, \cdots, I_{k_M}$. Meanwhile we have $\forall 1 \leq j \leq M$, $m(I_{k_j}) \leq \dfrac{1}{\varepsilon} \displaystyle\int_{I_{k_j}} |f|$ and thus

$$\begin{aligned} m(K) \leq 3 \sum_{j=1}^{M} m(I_{k_j}) &\leq \frac{3}{\varepsilon} \sum_{j=1}^{M} \int_{I_{k_j}} |f| \\ &= \frac{3}{\varepsilon} \int_{\bigcup_{j=1}^{M} I_{k_J}} |f| \\ &\leq \frac{3}{\varepsilon} \|f\|_1 \end{aligned}$$

■

Corollary

Corollary 16. *Given $f \in \mathscr{L}^1(\mathbb{R})$, then f^* is finite valued a.e.*

Proof. $\forall N > 0$, $m(\{f^* > N\}) \leq \frac{3}{N}\|f\|_1 \to 0$, as $N \to +\infty$, and $m(\{f^* = +\infty\})$ is a null set. ■

REMARK

$f \in \mathscr{L}^1(\mathbb{R})$ **DO NOT** imply $f^* \in \mathscr{L}^1(\mathbb{R})$ in general. To see this, suppose $f = \mathbf{1}_{[-1,1]} \in \mathscr{L}^1(\mathbb{R})$ and

$$f^* := \sup_{I \in \mathscr{I}(x)} \frac{1}{m(I)} \int_I |f|$$

in general,

$$\frac{1}{b-a}\int_a^b f(t)dt = \begin{cases} 0, \text{if } (a,b) \cap [-1,1] = \varnothing \\ \\ \frac{\min\{b,1\} - \max\{a,-1\}}{b-a}, \text{otherwise} \end{cases}$$

From here, we can derive that

$$f^*(x) = \begin{cases} 1, \text{if } x \in (-1,1) \\ \\ \frac{2}{|x|+1}, \text{otherwise} \end{cases}$$

and $f^*(x) \notin \mathscr{L}^1(\mathbb{R})$.

Lebesgue Differentiation Theorem

Theorem

Theorem 40. *Given $f \in \mathscr{L}^1(\mathbb{R})$, then for a.e $x \in \mathbb{R}$, if $\{I_n\}$ is a sequence of open intervals such that $x \in I_n$ for every n and $\lim_{n\to+\infty} m(I_n) = 0$ ($\{I_n\}$ is a sequence shrinking to x), then*

$$\lim_{n\to+\infty} \frac{1}{m(I_n)} \int_{I_n} |f(t) - f(x)| dt = 0$$

In particular,

$$\lim_{n\to+\infty} \frac{1}{m(I_n)} \int_{I_n} f(t) dt = f(x)$$

The last statement (in particular part) is due the fact that (assume x is such that $|f(x)| < +\infty$)

$$\left| \frac{1}{m(I_n)} \left(\int_{I_n} f(t) dt \right) - f(x) \right| = \left| \frac{1}{m(I_n)} \int_{I_n} (f(t) - f(x)) dt \right| \leq \frac{1}{m(I_n)} \int_{I_n} |f(t) - f(x)| dt$$

Proof. *W.L.O.G*, assume that f is finite valued **EVERYWHERE**, and we will only consider finite intervals I_n. For every $k \geq 1$, consider

$$B_k := \left\{ x \in \mathbb{R} : \exists \{I_n\} \subseteq \mathscr{I}(x) \text{ with } \lim m(I_n) = 0 \text{ and } \lim_{n\to+\infty} \frac{1}{m(I_n)} \int_{I_n} |f(t) - f(x)| dt \geq \frac{1}{k} \right\}$$

Thus $\{B_k\}$ is increasing, and

$$\bigcup_{k=1}^{+\infty} B_k := \{x \in \mathbb{R} : \text{the statement of the theorem fails}\}.$$

It's now sufficient to show that $m(B_k) = 0$ for every k. Fix an arbitrary $\varepsilon > 0$, since $C_C(\mathbb{R})$ is dense in $\mathscr{L}^1(\mathbb{R})$, thus $\exists g \in C_C(\mathbb{R})$ such that we have $\|f - g\|_1 \leq \varepsilon$. Since g is continuous and supported on a compact set, thus for all $k \geq 1$ and x, there exists $\alpha > 0$ such that

$$|t - x| \leq \alpha \implies |g(t) - g(x)| \leq \frac{1}{3k}$$

$\forall x$, given any sequence $\{I_n\} \subseteq \mathscr{I}(x)$ such that $\lim_{n\to\infty}(I_n) = 0$, we have

$$\frac{1}{m(I)} \int_{I_n} |f(t) - f(x)| dt \leq \underbrace{\frac{1}{m(I_n)} \int_{I_n} |f(t) - g(t)| dt}_{(*)} + \underbrace{\frac{1}{m(I_n)} \int_{I_n} |g(t) + g(x)| dt}_{(**)} + \underbrace{|g(x) - f(x)|}_{(***)}$$

According to the analysis of g, we know that when n is sufficiently large such that $m(I_n) \leq \alpha$, we have $|g(t) - g(x)| \leq \frac{1}{3k}, \forall t \in I_n$, so $(**) \leq \frac{1}{3k}$. In order for x to be in B_k, we need $\limsup((*) + (**) + (***)) \geq \frac{1}{k}$, since $(**) \leq \frac{1}{3k}$, we must have

$$\limsup((*) + (***)) \geq \frac{2k}{3}$$

Now we define

$$C_k := \left\{x \in \mathbb{R} : \limsup(*) \geq \frac{1}{2k}\right\} \text{ and } D_k := \left\{x \in \mathbb{R} : \limsup(***) \geq \frac{1}{3k}\right\}.$$

Thus $m(B_k) \leq m(C_k) + m(D_K)$ and

$$m(D_k) = m\left(\left\{|f-g| \geq \frac{1}{3k}\right\}\right) \leq sk\|f-g\|_1 \leq 3k\varepsilon$$

and

$$\begin{aligned} m(C_K) &= m\left(\left\{\limsup \frac{1}{m(I_n)}\int_{I_n}|f-g| \geq \frac{1}{3k}\right\}\right) \\ &\leq m\left(\left\{(f-g)^* \geq \frac{1}{3k}\right\}\right) \\ &\leq 3\cdot 3k\|f-g\|_1 \end{aligned}$$

Thus $m(B_k) \leq 12k\varepsilon$, $\forall \varepsilon > 0$. ■

REMARK

In the statement of the Lebesgue differentiation theorem, a.e $x \in \mathbb{R}$ can not be replaced by point-wise. To see this, construct $f = \mathbf{1}_{[0,1]} \in \mathscr{L}^1(\mathbb{R})$, and $I_n = \left(-\frac{1}{n}, \frac{1}{n}\right) \in \mathscr{I}(0)$, then

$$\frac{1}{m(I_n)}\int_{I_n} f(t)dt = \frac{2}{n}\int_{-\frac{1}{n}}^{\frac{1}{n}} f(t)dt = \frac{1}{2} \neq 0.$$

and thus 0 is not a Lebesgue point.

Corollary

Corollary 17. *If $f \in \mathscr{L}^1(\mathbb{R})$, then a.e $x \in \mathbb{R}$,*

$$\lim_{n\to+\infty} \frac{1}{2h}\int_{x-h}^{x+h} \left|f(t) - f(x)\right| dt = 0$$

and equivalently,

$$\lim_{h\to 0} \frac{1}{2h}\int_{-h}^{h} \left|f(x+y) - f(x)\right| dy = 0$$

Proof. Note that it is trivial that

$$\lim_{n\to+\infty} \frac{1}{2h}\int_{x-h}^{x+h} \left|f(t) - f(x)\right| dt = 0 \implies \lim_{h\to 0} \frac{1}{2h}\int_{-h}^{h} \left|f(x+y) - f(x)\right| dy = 0$$

If $x \in \mathbb{R}$ is such that the original equation will fail, then there exists $\{h_n\} \in \mathbb{R}^+$ such that $\lim_{n\to+\infty} h_n = 0$ and

$$\lim_{h\to+\infty} \frac{1}{2}\int_{x-h_n}^{x+h_n} |f(t) - f(x)| dt \neq 0$$

thus the Lebesgue differentiation theorem fails at x with $I_n = (x - h_n, x + h_n)$, thus x is not a Lebesgue point of f, and so that $\{x \in \mathbb{R} : \text{the equation fails}\}$ is a null set.

■

Corollary

Corollary 18. *Given $f \in \mathscr{L}^1(\mathbb{R})$, f^* be the HL-maximal function of f, then $|f| \leq f^*$ a.e.*

Proof. Apply Lebesgue differentiation theorem to $|f|$, thus for a.e $x \in \mathbb{R}$ with $I_n = \left(x - \frac{1}{n}, x + \frac{1}{n}\right)$, we have

$$|f(x)| = \lim_{n \to +\infty} \frac{1}{m(I_n)} \int_{I_n} |f(t)| dt \leq \sup \left\{ \frac{1}{m(I)} \int_I |f| \right\} \stackrel{\text{def}}{=} f^*(x).$$

■

Corollary

Corollary 19. *If $f \in \mathscr{L}^1(\mathbb{R})$, then $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $\forall$ interval I with $m(I) \leq \delta$, we have*

$$\int_I |f| \leq \varepsilon.$$

In fact, I can be replace by any $A \in \mathscr{M}$ measurable.

We will prove this corollary based on the fact that I is an interval, but it can be generalized for all measurable sets as the theorem suggested.

Proof. Recall that $f^* < +\infty$ a.e, if $A_N := \{x \in \mathbb{R} : f^*(x) > N\}$, then $\mathbf{1}_{A_N} \to 0$ as $N \to +\infty$, thus

$$\lim_{N \to +\infty} \int_{\mathbb{R}} \mathbf{1}_{A_N} |f| = 0$$

Then by DOM, we have

$$\lim_{N \to +\infty} \int_{f^* > N} |f| = 0$$

Then $\forall \varepsilon > 0, \exists N > 0$ such that

$$\int_{f^* > N} |f| < \frac{\varepsilon}{2}$$

Now with this choice of N, set $\delta = \frac{\varepsilon}{2N}$, then *W.L.O.G* assume that I is an open interval, then choose $m(I) \leq \delta$ and we then have

$$\begin{aligned} \int_I |f| &= \int_{I \cap \{f^* > N\}} |f| + \int_{I \cap \{f^* \leq N\}} |f| \\ &\leq \underbrace{\int_{\{f^* > N\}} |f|}_{\leq \varepsilon/2} + \underbrace{N \cdot m(I)}_{\leq N\delta\varepsilon/2} \leq \varepsilon \end{aligned}$$

■

Theorem

Theorem 41. *Given $f \in \mathscr{L}^1(\mathbb{R})$ and $a \in \mathbb{R}$, define*

$$F(x) = \int_a^x f(t)dt \quad \text{for} \quad x \geq a$$

Then F is uniformly continuous and $F'(x)$ exists and $F'(x) = f(x)$ for a.e $x \in \mathbb{R}$.

Proof. We want to show that

$$F(x) = \int_a^x f(t)dt$$

is uniformly continuous, so $\forall \varepsilon > 0$ let δ be constructed in the previous corollary, then $\forall x > y > a$ such that $|x - y| \leq \delta$, we have

$$\| F(x) - F(y)| = \left| \int_y^x f(t)dt \right| \leq \int_{(y,x)} |f(t)|dt \leq \varepsilon.$$

and thus F is uniformly continuous.

Next, we will show that F' exists a.e, to show that $F' = f$ a.e and

$$\lim_{h \to 0} \frac{1}{h}(F(x+h) - F(x)) = f(x)$$
$$\lim_{h \to 0} \frac{1}{h}(F(x) - F(x-h)) = f(x)$$

for a,e $x \in \mathbb{R}$.

If $x \in \mathbb{R}$ is a Lebesgue point of f and $f(x)$ is finite-valued, then we have

$$\begin{aligned}
\left| \frac{1}{h}(F(x+h) - F(x)) - f(x) \right| &= \left| \frac{1}{h} \int_x^{x+h} f(t)dt - f(x) \right| \\
&= \left| \frac{1}{h} \int_x^{x+h} [f(t) - f(x)]dt \right| \\
&\leq \frac{1}{h} \int_x^{x+h} |f(t) - f(x)|dt \\
&\leq 2 \cdot \frac{1}{2h} \int_{x-h}^{x+h} |f(t) - f(x)|dt \xrightarrow{h \to 0} 0.
\end{aligned}$$

And similar statements can be drawn for the other equation.

■

REMARK

In general, a.e $x \in \mathbb{R}$ cannot be dropped, to see this, construct $f = \mathbf{1}_{\{0\}}$, then

$$F = \int_a^x f, F \equiv 0, F'(0) \neq f'(0)$$

Increasing Functions

Let F be an increasing function on $[a,b]$ (we restrict to finite valued functions). If needed, whenever necessary, we can extend $F(x)$ to $\mathbb{R}$ by $F(x) = F(a)$ for all $x < a$ and $F(x) = F(b)$ for all $x > b$.

Proposition

Proposition 68. *Assume that F is increasing on $[a,b]$, then f is continuous except at countably many points in $[a,b]$.*

Proof. Already proven in MATH 255. ∎

For a general function $x \in \mathbb{R}$, we define the followings:

$$\overline{D_r}F(x) = \limsup_{h\to 0} \frac{F(x+h)-F(x)}{h}; \quad \underline{D_r}F(x) = \liminf_{h\to 0} \frac{F(x+h)-F(x)}{h}$$

$$\overline{D_l}F(x) = \limsup_{h\to 0} \frac{F(x)-F(x-h)}{h}; \quad \underline{D_l}F(x) = \liminf_{h\to 0} \frac{F(x)-F(x-h)}{h}$$

If all 4 limits match, then F' exists and is equal to any of the 4 listed.

Theorem

Theorem 42. *(Vitali's Covering Theorem)*

Given a set $E \subseteq \mathbb{R}$, a collection $\mathscr{J}$ of intervals is called a Vitali covering of E if $\forall x \in E$, $\forall \varepsilon > 0$, $\exists I \in \mathscr{J}$, such that $x \in I, m(I) < \varepsilon$. If $m(E) < +\infty$, $\mathscr{J}$ is a Vitalli covering of E, then $\forall \varepsilon > 0$, $\exists I_1, I_2, \cdots, I_N \in \mathscr{J}$ such that $I_i \bigcap I_j = \varnothing, i \neq j$ and

$$m\left(E \Big\backslash \bigcup_{n=1}^{N} I_n\right) \leq \varepsilon.$$

Proof. *W.L.O.G*, assume that $m\left(\bigcup_{I\in\mathscr{J}} I\right) < +\infty$, now define

$$\alpha_1 = \sup\{m(I) : I \in \mathscr{J}\} < +\infty$$

then $\exists I_1 \in \mathscr{J}$, such that $m(I_1) > \frac{\alpha}{2}$, set

$$\mathscr{J}_1 = \{I \in \mathscr{J}, I \cap I_1 = \varnothing\}$$

and define

$$\alpha_2 = \sup\{m(I) : I \in \mathscr{J}_1\} < +\infty$$

so $\exists I_2 \in \mathscr{J}$, such that $m(I_2) > \frac{\alpha_2}{2}$. Continue this approach, define

$$\mathscr{J}_2 := \{I \in \mathscr{J}, I \cap I_1 = \varnothing, I \cap I_2 = \varnothing\}$$

and

$$\alpha_3 = \sup\{m(I) : I \in \mathscr{J}_2\}$$

Repeat this process, we'll get $\{\alpha_k\}, \{I_K\}$ such that $I_i \cap I_j = \varnothing, i \neq j$, and

$$\alpha_k := \sup\{m(I) : I \in \mathscr{J}, I \cap I_j = \varnothing, \forall j = 1, 2, \cdots, k-1\}.$$

Since we have

$$\bigcup_{k=1}^{+\infty} I_k \subseteq \bigcup_{I \in \mathscr{J}} I, \text{and } m\left(\bigcup_{k=1}^{+\infty} I_k\right) < +\infty,$$

thus by countable additivity,

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

since

$$\bigcup_{k=1}^{+\infty} I_k \subseteq \bigcup_{I \in \mathscr{J}} I, \sum_{k=1}^{+\infty} \frac{\alpha_k}{2} \leq \sum m(I_k) < +\infty, \alpha_k \to 0, \text{as } k \to +\infty.$$

Then fix any $\varepsilon > 0$, $\exists N$ large enough such that

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

CLAIM: $m\left(E \Big\backslash \bigcup_{i=1}^{N} I_i\right) < \varepsilon.$

To see this, note that it suffices to show that

$$m\left(E \Big\backslash \bigcup_{i=1}^{N} \bar{I}_i\right) < \varepsilon$$

Since $\bar{I}_i$ is closed and $\forall x \in E \backslash \bigcup_{i=1}^{N} \bar{I}_i$, we have

$$d\left(x, \bigcup_{i=1}^{N} \bar{I}_i\right) = \Lambda > 0$$

Recall that $\mathscr{J}$ is a Vitali covering of E, then $\exists I^* \in \mathscr{J}$ such that $x \in I^*$ and $m(I^*) < \Lambda$, thus it must be that

$$I^* \cap I_i = \varnothing, \forall i - 1, 2, \cdots, N$$

Thus $m(I^*) \leq \alpha_{N+1}$. Assume that $N^* \geq N+1$ such that $m(I^*) \leq \alpha_{N^*}$, then $\exists j = N+1, N^*$ such that $I^* \cap I_j \neq \varnothing$, and $m(I^*) \leq \alpha_{N^*} \leq \alpha_j \leq 2m(I_j)$, so $I^* \subseteq 5I_j$, and then

$$E \Big\backslash \bigcup_{i=1}^{N} \bar{I}_i \subseteq \bigcup_{k=N+1}^{+\infty} 5I_j$$

thus
$$m\left(E\Big\backslash\bigcup_{i=1}^{N}\bar{I}_i\right)\leq 5\sum_{k=N+1}^{+\infty}m(I_k)\leq\varepsilon$$

■

Proposition

Proposition 69. *Assume that F is an increasing function on $[a,b]$, then F' exists a.e on $[a,b]$.*

Proof. Clearly, $\forall x\in[a,b]$,
$$\overline{D_r}F(x)\geq\underline{D_r}F(x);\underline{D_l}F(x)\leq\underline{D_l}F(x)$$
So if we could just show
$$\overline{D_r}F(x)\leq\underline{D_l}F(x),\overline{D_l}F(x)\leq\underline{D_r}F(x)$$
then we're done (cooked).

We first will show that
$$\overline{D_l}F(x)\leq\underline{D_r}F(x)$$
Set $E=\{x\in\mathbb{R}:\overline{D_l}F(x)>\underline{D_r}F(x)\}$, we will show E is a null set. or $p,q\in\mathbb{Q}$ and $p<q$, set
$$E_{p,q}=\{x\in[a,b]:\overline{D_r}F(x)>q,\underline{D_r}F(x)<p\}$$
Then observe that
$$E=\bigcup_{p,q\in\mathbb{Q},p<q}E_{p,q},\text{and it is countable}$$
So we will show for each $E_{p,q}$, it's a null set. Now we will use Vitali's covering theorem to prove that $m(E_{p,q})=0$. Assume otherwise, then
$$m(E_{p,q})=\delta>0$$
Fix any $0<\varepsilon<\frac{\delta}{2}$, choose an open set $E_{p,q}\subseteq G$ such that
$$m(G)\leq m(E_{p,q})+\varepsilon=\delta+\varepsilon$$
Consider
$$\mathscr{J}:=\left\{I=[x,x+h]\subseteq G,\frac{F(x+h)-F(x)}{h}<p\right\}$$
Then for every $x\in E_{p,q}$, $x\in G$ open, we have
$$\underline{D_r}F(x)<p,\text{ie}\quad\liminf_{h\to 0}\frac{F(x+h)-F(x)}{h}<p$$
That means $\exists I=[x,x+h]\in\mathscr{J}$, for arbitrarily small h, then $\mathscr{J}$ is a Vitali covering of $E_{p,q}$. Then there exist disjoint sets $I_1,I_2,\cdots,I_N\in\mathscr{J}$, such that
$$m\left(E_{p,q}\Big\backslash\bigcup_{n=1}^{N}I_n\right)\leq\varepsilon$$

Assume $I_i = [x_i, x_i + h_i], i = 1, 2, \cdots, N$ and define

$$\widetilde{G} := \bigcup_{n=1}^{N} (x_i, x_i + h_i); \widetilde{E_{p,q}} := E_{p,q} \cap \widetilde{G}$$

thus we have

$$\begin{aligned} m(\widetilde{E_{p,q}}) &= m(E_{p,q} \cap \widetilde{G}) \\ &= m(E_{p,q} \cap (\bigcup I_n)) \\ &= m(E_{p,q}) - m\left(E_{p,q} \Big\backslash \bigcup_{n=1}^{N} I_n\right) \geq \delta - \varepsilon \end{aligned}$$

So since $\widetilde{E}_{p,q} \subseteq \widetilde{G}$ and G is open, then define

$$\widetilde{\mathscr{I}} := \left\{ \widetilde{I} = [y - r, y] \subseteq \widetilde{G}, \frac{F(y) - F(y - r)}{r} > q \right\}$$

Following the same argument, we know that $\widetilde{\mathscr{I}}$ is a Vitali covering of $\widetilde{E}_{p,q}$, so there are again disjoint sets $\widetilde{I}_1, \cdots, \widetilde{I}_M \in \widetilde{\mathscr{I}}$ such that

$$m\left(\widetilde{E}_{p,q} \Big\backslash \bigcup_{j=1}^{M} \widetilde{I}_j\right) \leq \varepsilon$$

and

$$m\left(\bigcup_{j=1}^{M} \widetilde{I}_j\right) = m(\widetilde{E}_{p,q}) - m\left(\widetilde{E}_{p,q} \Big\backslash \bigcup_{j=1}^{M} \widetilde{I}_j\right) geq \delta - 2\varepsilon$$

Now assume $\widetilde{I}_j := [y_j - r - j, y_j]$ then we know that those are sub-intervals of $I_1, I_2, \cdots, I_n$ respectively. Since F is increasing, it means that

$$\underbrace{\sum_{j=1}^{M} (F(y_j) - F(y_j - r_j))}_{\text{L.H.S}} \leq \underbrace{\sum_{j=1}^{N} (F(x_i + h_i) - F(x_i))}_{\text{R.H.S}}$$

and

$$R.H.S \leq p \sum h_i = p \cdot m\left(\bigcup_{i=1}^{N} I_i\right) \leq p \cdot m(G) \leq p(\delta + \varepsilon)$$

$$L.H.S \geq q \sum r_j = q \cdot m\left(\bigcup_{j=1}^{M} I_j\right) \geq q(\delta - 2\varepsilon)$$

which implies that

$$q(\delta - 2\varepsilon) \leq p(\varepsilon + \delta)$$

and thus

$$q\delta \leq p\delta,$$

a contradiction.

The same idea may be applied for showing $\overline{D_r}F(x) \leq \underline{D_l}F(x)$, and that finishes the proof.

■

Proposition

Proposition 70. *Assume $F : [a,b] \to \mathbb{R}$ is increasing, then $F' \in \mathscr{L}([a,b])$ and $F'(x) > 0$ a.e on $[a,b]$ and*

$$\int_a^b F'(t)dt \leq F(b) - F(a)$$

Proof. Clearly, $F'(x) \geq 0$ a.e on $[a,b]$ follows from the previous proposition. For a.e $x \in [a,b]$, we have

$$F'(x) = \lim_{n\to+\infty} G_n(x) \;\text{ where }\; G_n(x) = \frac{F\left(x+\frac{1}{n}\right) - F(x)}{\frac{1}{n}}$$

We can expand F to $\mathbb{R}$ if needed (followed by the previous construction), and by Fatou's lemma,

$$\int_a^b F'(x)dx \leq \liminf_n \int_a^b G_n(x)dx$$

and

$$\begin{aligned}
\lim_{n\to+\infty} G_n(x) &= \liminf_n \left(n \left(\int_a^b F\left(x + \frac{1}{n}\right) dx - \int_a^b F(x)dx \right) \right) \\
&= \liminf_n \left(n \left(\int_{a+1/n}^{b+1/n} F(t)dt - \int_a^b F(t)dt \right) \right) \\
&= \liminf_n \left(n \left(\int_b^{b+1/n} F(t)dt - \int_a^{a+1/n} F(t)dt \right) \right) \\
&\leq \liminf_n \left(n \left(F(b) \cdot \frac{1}{n} - F(a) \cdot \frac{1}{n} \right) \right) \\
&= F(b) - F(a).
\end{aligned}$$

■

REMARK

The inequality $\int_a^b F'(t)dt \leq F(b) - F(a)$ can be a strict inequality. Think of $f : [0,1] \to \mathbb{R}$ to be the Cantor-Lebesgue function, then we know that $f'(x) \equiv 0$ a.e on $[0,1]$, since $\forall x \in [0,1] \backslash \mathscr{C}$, $f(x)$ is just constant, and such set has full measure. So

$$\int_0^1 f'(t)dt = 0 < f(1) - f(0) = 1$$

Functions With Bounded Variation

Definition

Definition 35. *We define a function $F : [a,b] \to \mathbb{R}$ is of bounded variation on $[a,b]$, denoted by $F \in BV([a,b])$ if*

$$T_F(a,b) := \sup\left\{\sum_{k=1}^{N} |F(x_k) - F(x_{k-1})| \;\; : \;\; N \geq 1, a = x_1 < x_2 < \cdots < x_N = b\right\} < +\infty$$

$T_F(a,b)$ is called the total variation of F over $[a,b]$.

Some properties follow:

Properties of Bounded Variation

(i) For $x \in [a,b]$, set $V_F(x) = T_F(a,x)$, then $V_F(x)$ is an increasing function and is always non-negative, and $\forall a \leq x < y \leq b, V_F(y) - V_F(x) = T_F(x,y)$, i.e

$$T_F(a,x) + T_F(x,y) = T_F(a,y)$$

(ii) If $F, G \in BV([a,b])$, then $F \pm G \in BV([a,b])$ and

$$T_{F+G}(a,b) \leq T_F(a,b) + T_G(a,b)$$

(iii) If F is monotonic, then $F \in BV([a,b])$.
• We may just assume F is increasing, then $\forall$ partition $a = x_0 < x_1 < \cdots < x_N = n$, we have

$$\sum_{k=1}^{N} |F(x_k) - F(x_{k-1})| = \sum_{k=1}^{N} F(x_k) - F(x_{k-1}) = F(b) - F(a)$$

and $T_F(a,b) = |F(b) - F(a)|$ for general monotonic functions.
(iv) If $f \in \mathscr{L}^1(\mathbb{R})$ and $F(x) = \displaystyle\int_a^x f(t)dt$ or $x \in [a,b]$, then $F \in BV([a,b])$ and

$$T_F(a,b) \leq \|f\|_1 = \int_a^b |f(t)|dt.$$

• $\forall a = x_0 < x_1 < \cdots < x_N = b$, we have

$$\sum_{k=1}^{N} |F(x_k) - F(x_{k-1})| = \sum_{k-1}^{N} \left| \int_{x_{k-1}}^{x_k} f(t)dt \right| \leq \sum_{k=1}^{N} \int_{x_{k-1}}^{x_k} |f(t)|dt = \int_a^b |f(t)|dt < +\infty$$

(v) If $F \in BV([a,b])$, then F is bounded on $[a,b]$.

• $\forall x \in [a,b]$, we have

$$\begin{aligned} |F(x)| &\leq |F(x)-F(a)| + |F(a)| \\ &\leq T_F(a,x) + |F(a)| \\ &\leq T_F(a,b) + |F(a)| < +\infty \end{aligned}$$

(vi) If F is continuous on $[a,b]$ and differentiable everywhere on (a,b) and $\exists M > 0$ such that $|F'(x)| \leq M$, then $\forall x \in (a,b)$, then $F \in BV([a,b])$.
• By the mean value theorem, $\forall x,y \in (a,b), x<y$, $\exists z \in (a,y)$, such that

$$\frac{F(y)-F(x)}{y-x} = F'(z)$$

For every such y, we then have $|F(y)-F(x)| \leq M(y-x)$ and so $\forall a = x_1 < x_2 < \cdots < x_n = b$, we have

$$\sum_{k=1}^{N} |F(x_k) - F(x_{k-1})| \leq M(b-a) < +\infty$$

(vii) In proposition (vi), the assumption $|F'| \leq M$ **cannot** be dropped.

Theorem

Theorem 43. *A function $F : [a,b] \to \mathbb{R}$ is in $BV([a,b])$ if and only if there exist $H, G : [a,b] \to \mathbb{R}$ such that H, G are all increasing and $F(x) = H(x) - G(x)$ for all $x \in [a,b]$.*

Proof. ($\Longleftarrow$) :If $F(x) = H(x) - G(x)$ with H,G both increasing, and $H,G \in BV([a,b])$, then by previous properties, we know that this direction is trivial.
($\Longrightarrow$) : Assume $F \in BV([a,b])$, define $H(x) = T_F(a,x)$ for all $x \in [a,b]$, and we claim that $H(x)$ is increasing by definition. Also let $G(x) = H(x) - F(x)$, then $\forall x,y \in [a,b]$ with $x<y$,

$$\begin{aligned} G(y) - G(x) &= H(y) - H(x) - (F(y) - F(x)) \\ &= T_F(x,y) - (F(y) - F(x)) \\ &\geq T_F(x,y) - |F(y) - F(x)| \geq 0. \end{aligned}$$

and thus $G(x)$ is increasing.

■

Theorem

Theorem 44. *If $F : [a,b] \to \mathbb{R}$ is in $BV([a,b])$, then F is continuous on $[a,b]$ except for at most countably many points; F' exists a.e on $[a,b]$; $F' \in \mathscr{L}^1([a,b])$.*

Absolutely Continuous Functions

Definition

Definition 36. *Given $F : [a,b] \to \mathbb{R}$ is called absolutely continuous on $[a,b]$, denoted by $F \in AC([a,b])$, if $\forall \varepsilon > 0$, $\exists \delta > 0$, such that for any disjoint intervals $(a_k, b_k) \subseteq (a,b)$, $k = 1, 2, \cdots, N$ with $\sum_{k=1}^{N} |a_k - b_k| \leq \delta$, we have*

$$\sum_{k=1}^{N} |F(b_k) - F(a_k)| \leq \varepsilon.$$

REMARK

Note that (a_k, b_k)'s are not necessarily a partition of (a,b).

Proposition

Proposition 71. *If $F \in AC([a,b])$, then F is uniformly continuous on $[a,b]$.*

Proof. $\forall \varepsilon > 0$, let $\delta > 0$, $\forall x, y \in [a,b]$, if $|y - x| < \delta$, then $|F(y) - F(x)| < \varepsilon$. ■

Proposition

Proposition 72. *If $f \in AC([a,b])$ and $F(x) = \int_a^x f(t)dt$, $t \in [a,b]$, then F is absolutely continuous.*

Proof. Recall that $\forall \varepsilon > 0$, $\exists M$, such that

$$\int_{\{f^* > M\}} |f| < \frac{\varepsilon}{2}.$$

then set $\delta = \varepsilon/2M$, then for all $(a_k, b_k) \subseteq (a,b)$ disjoint, where $k = 1, 2, \cdots, N$, such that $\sum_{k=1}^{N} |b_k - a_k| < \delta$,

$$\begin{aligned}
\sum_{k=1}^{N} |F(b_k) - F(a_k)| &< \sum_{k=1}^{N} \int_{a_k}^{b_k} |f| \\
&= \sum_{k=1}^{N} \left[\int_{(a_k,b_k) \cap \{f^* > M\}} |f| + \int_{(a_k,b_k) \cap \{f^* \leq M\}} |f| \right] \\
&\leq \int_{\bigcup_{k=1}^{N} (a_k,b_k) \cap \{f^* > M\}} |f| + M \sum_{k=1}^{N} |b_k - a_k| \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.
\end{aligned}$$

■

Proposition

Proposition 73. *If $F, G \in AC([a,b])$, then trivially $F \pm G$ are both absolutely continuous.*

Proof. :) ■

Proposition

Proposition 74. *If F is absolutely continuous, then $F \in BV([a,b])$.*

Proof. Set $\varepsilon = 1$, take $\delta > 0$, consider a partition of (a,b):

$$a = t_0 < t_1 < \cdots < t_L = b$$

such that

$$t_{i+1} = t_i + \frac{b-a}{L} \quad , \frac{b-a}{L} \leq \delta$$

Now for each $i = 0, 1, \cdots, L$, for any partition of $[t_i, t_{i+1}]$, where

$$t_i = x_0 < x_1 < \cdots < x_N = t_{i+1}$$

and $(x_k . x_{k+1})$ are disjoint, and

$$\sum_{k=0}^{N-1} |x_{k+1} - X_k| - x_k| = |t_{i+1} - t_i| \leq \delta$$

which implies

$$\sum_{k=0}^{N-1} |F(x_{k+1} - x_k)| \leq 1 \implies T_F(t_i, t_{i+1}) \leq 1$$

Then $T_F(a,b) = \sum_{i=0}^{L-1} T_F(t_i, t_{i+1}) \leq L < +\infty$ ■

Proposition

Proposition 75. *If F is continuous on $[a,b]$ and differentiable everywhere on (a,b) and $|F'(x)| \leq M$, (where M is a constant) for all $x \in (a,b)$, then $F \in AC([a,b])$.*

Proof. We used the mean value Theorem to show that

$$\forall x, y \in [a,b], |F(x) - F(y)| \leq M|x-y|.$$

■

Theorem

Theorem 45. *If $F \in AC([a,b])$, then F' exists a.e. on $[a,b]$ and furthermore $F' \in \mathscr{L}^1([a,b])$.*

Proof. The following is immediately apparent from the fact that the *AC* functions are a subset of the *BV* functions. ■

Theorem

Theorem 46. *Given $F \in AC([a,b])$, F is a constant on $[a,b]$ if and only if $F'(x) = 0$ a,e. on $[a,b]$.*

Proof. Trivial. ∎

REMARK

The condition $F \in AC([a,b])$ cannot be dropped, think of F being the Cantor-Lebesgue function.

Theorem

Theorem 47. *(Fundamental Theorem of Calculus)*
If $F \in AC([a,b])$ then F' exists a.e. on (a,b), $F' \in \mathscr{L}^1([a,b])$ and $\forall x \in [a,b]$, we have

$$F(x) - F(a) = \int_a^x F'(t)dt$$

Proof. Assume $F \in AC([a,b])$, then it implies F' exists a.e. on (a,b). $F' \in \mathscr{L}^1([a,b])$, define

$$G(x) = F(a) + \int_a^x F'(t)dt, \forall x \in [a,b]$$

Then according to property 72, $G(x) \in AC([a,b])$, and by the theorem 41, we know that $G'(x) = F'(x)$ a.e on $[a,b]$, that means take $H = F - G \in AC([a,b])$ and $H' = F' = G' = 0$, thus H is a constant, thus

$$H(x) = H(a) = 0, \forall x \in [a,b]$$

Then

$$F(x) = G(x) = F(a) + \int_a^x F'(t)dt, \forall x \in [a,b].$$

∎

What we have covered so far is basically the end of this course (MATH 454 Honors Analysis 3). But there are way more topics to learn in analysis. One of the most important applications to the topics we have covered so far is probability theory. Probability is a special measure with total measure being 1, and many topics in probability theory have direct connection with measure theory and Lebesgue integration. Furthermore, in MATH 455 (Honors Analysis 4), we will focus on more abstract spaces like Hilbert spaces and functional analysis. I hope you will like it as well :).

Finally, I would like to express my deepest gratitude to my professor who taught this course amazingly: Prof. Linan Chen.

Chapter 5

What's Next?

Based on what we have established so far, we would like to move a little bit forward to probability measure (a special measure we introduced with ”total mass” being 1).

Definition

Definition 37. *Assume μ is a probability measure on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, define a function F_μ by $F_\mu(x) = \mu((-\infty, x]), \forall x \in \mathbb{R}$. Then we call F_μ to be the distribution function of μ.*

Furthermore, if X is a random variable on $(\Omega, \mathscr{F}, \mathbb{P})$, then

$$\mu((-\infty, x]) = \mathbb{P}(X \leq x)$$

Based on the definition, we have the following properties of F_μ:

PROPERTIES OF A DISTRIBUTION FUNCTION

(i) F_μ is always increasing;

(ii) $\lim_{x \to +\infty} F_\mu(x) = 1$ and $\lim_{x \to -\infty} F_\mu(x) = 0$;

(iii) F_μ has at most countably many discontinuities and F_μ is *R.C.L.L* (right continuous with left hand limit);

(iv) $\forall x \in \mathbb{R}$, $F_\mu(x^+) = F_\mu(x)$ and $F_\mu(x^-)$ exists;

(v) $\forall a, b \in \mathbb{R}$ with $a < b$, $F_\mu(b) - F_\mu(a) = \mu((a, b])$;

(vi) The distribution function uniquely determines μ;

(vii) In fact, any $F : \mathbb{R} \to [0, 1]$ that satisfies (i), (ii), (iii) is the distribution function of some probability measure.

Definition

Definition 38. *Let μ be a probability measure on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ with distribution function F_μ, we say μ is absolutely continuous with Lebesgue measure, denoted by $\mu << m$, if F_μ is AC on $\mathbb{R}$, i.e $\exists f \in \mathscr{L}^1(\mathbb{R})$, such that $\forall a \in \mathbb{R}$, $\forall x > a$, we have*

$$F_\mu(x) - F_\mu(a) = \int_a^x f(t)dt.$$

and f is called the probability density function of μ.

Proposition

Proposition 76. *If μ is absolutely continuous with Lebesgue measure with probability density function f, then $\forall B \in \mathfrak{B}(\mathbb{R})$, we have*

$$\mu(B) = \int_B f(t)dt.$$

Corollary

Corollary 20. *Assume $\mu << m$ with density function f, then if $g : \mathbb{R} \to \mathbb{R}$ is $\mathfrak{B}(\mathbb{R})$-measurable, then*

$$\underbrace{\int_{\mathbb{R}} |g| d\mu}_{\text{the integral of } |g| \text{ w.r.t } \mu} = \underbrace{\int_{\mathbb{R}} |g(x)| f(x) dx}_{\text{Lebesgue integral}}.$$

In particular, this corollary implies that $g \in \mathscr{L}^1(\mu)$ if and only if $g \cdot f \in \mathscr{L}^1(\mathbb{R})$.

Theorem

Theorem 48. *(Radon-Nikodym)*

Let μ, ν to be two σ-finite measure defined on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, if $\forall A \in \mathfrak{B}(\mathbb{R})$, $\nu(A) = 0 \implies \mu(A) = 0$, i.e ν-null sets are a subset of μ-null sets, then $\mu << \nu$ and hence $\exists f \in \mathscr{L}^1(\nu)$, such that $\forall B \in \mathfrak{B}(\mathbb{R})$,

$$\mu(B) = \int_B f d\nu,$$

we call f to be the Radon-Nikodym derivative of μ w.r.t ν, denoted by

$$f = \frac{d\mu}{d\nu}.$$

Theorem

Theorem 49. *(Lebesgue Decomposition Theorem)*

Given μ *to be any probability measure defined on* $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, μ *admits a unique decomposition :* $\mu = \mu_a + \mu_s$ *where* μ_a, μ_s *are also finite measures, such that*

(i) μ_a *is a finite measure on* $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ *and* $\mu_a << m$*;*

(ii) μ_s *is a finite measure on* $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ *such that* $\mu_s \perp m$*, i.e* μ_s *is singular to m :* $\exists E \in \mathfrak{B}(\mathbb{R}))$ *such that* $m(E) = 0$ *but* $\mu_s(E^C) = 0$*.*

Chapter 6

References

---

1. MATH 454 course notes written by Prof. Linan Chen, McGill University.

2. Stein Real Analysis : measure theory, integration and Hilbert spaces.

3. Measure, integration and real analysis by Sheldon Axler.

4.