Summer Workshop in Measure Theory

Ching-Cheong Lee

Contents

Part I

Summer 2010-2011

Chapter 1

Metric Space

The purpose of this chapter is to introduce basic concept and terminology in point-set topology used in this text and also in other branches of mathematics in which the language of topology is used. Although here we only attach our focus on metric space, most of the definition can be easily translated to a more general concept called topological space.

Throughout our text the notation $\mathbb{K}$ will mean either $\mathbb{R}$ or $\mathbb{C}$. By normed space we mean normed vector space.

1.1 Metric

Definition 1.1.1. A **metric space** is a nonempty set M with a function $d : M \times M \to \mathbb{R}$ such that for every $x, y, z \in M$,

(i) $d(x,y) \geq 0$ and equality holds iff $x = y$;

(ii) $d(x,y) = d(y,x)$;

(iii) (Triangle Inequality) $d(x,z) \leq d(x,y) + d(y,z)$.

Such a function d is called **metric**. Sometimes we may write (M,d) to denote a space M endowed with a metric d. Henceforth we write M instead of (M,d) and d is always the metric on the space mentioned unless there are two spaces.

Example 1.1.2. On any set X, the function d_{dis} defined by

$$d_{\text{dis}}(x,y) = \begin{cases} 0, & x = y, \\ 1, & x \neq y \end{cases}$$

is a metric, called discrete metric. And (X, d_{dis}) is called a discrete metric space. □

Example 1.1.3. On $C[0,1] := \{f : f \text{ is continuous on } [0,1]\}$, let $f, g \in C[0,1]$, then for $p \geq 1$,

$$d_p(f,g) := \sqrt[p]{\int_0^1 |f(x) - g(x)|^p \, dx} \quad \text{and} \quad d_\infty(f,g) := \sup_{x \in [0,1]} |f(x) - g(x)|$$

are metrics on $C[0,1]$. The verification of d_p being a metric follows from the celebrated Minkowski inequality, and that of d_∞ is easy. In fact, $(C[0,1],d_\infty)$ is a **complete metric space** which we may discuss in the future (roughly speaking, a space is said to be complete if any Cauchy sequence has a limit in that space). It can be checked that in Riemann integral, the following holds for $f \in C[0,1]$,

$$\lim_{n\to\infty}\left(\int_0^1 |f(x)|^n\,dx\right)^{1/n} = \sup_{x\in[0,1]} |f(x)|,$$

this inspires the definition d_∞. □

Example 1.1.4. All normed vector spaces are in particular a metric space. Recall that $\|\cdot\|$ on a vector space V satisfies the following:

(i) $\|x\| \ge 0$ and $\|x\| = 0 \iff x = 0$; (Positivity)

(ii) $\|\alpha x\| = |\alpha|\|x\|$ for $\alpha \in \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$); (Scaling Property)

(iii) $\|x+y\| \le \|x\| + \|y\|$. (Triangle Inequality)

It can be seen that $\|\cdot\|$ is a metric. Since $C[0,1]$ is a vector space, $(C[0,1],d_\infty)$ is in fact a complete normed vector space which is also called **Banach space**. □

Example 1.1.5. Let X,Y be two normed vector spaces, define $L(X,Y)$ to be the collection of all continuous linear transformations from X to Y (assuming the open subsets of X and Y are generated by "ball"s). $L(X,Y)$, by definition, is a vector space and on which we can define a norm by, letting $T \in L(X,Y)$,

$$\|T\| = \sup\{\|Tx\| : x \in X, \|x\| = 1\}.$$ □

We have had a few examples of metric space. The last 2 examples are intensively studied in MATH371. You will be asked to verify the norms defined above are really a norm in presentation.

1.2 Ball, Open Subsets

We define a ball $B(a,\epsilon) = \{x \in M : d(x,a) < \epsilon\}$ on a metric space M. Different choices of metric would induce different kind of balls, hence if necessary to distinguish two balls induced by d_1 and d_2, we may write $B_{d_1}(a,\epsilon)$ and $B_{d_2}(a,\epsilon)$ respectively.

Definition 1.2.1. A subset U of a metric space M is said to be open in M if

$$u \in U \implies B(u,\delta) \subseteq U, \text{ for some } \delta > 0.$$

Example 1.2.2. $(0,1)$ is open in $(\mathbb{R},|\cdot|)$, since when $x \in (0,1)$, $B(x,\min\{x,1-x\}) \subseteq (0,1)$. On (M,d_{dis}), any subset of M is open in M since $x \in M \implies B_{d_{\text{dis}}}(x,1) = \{x\}$. □

Proposition 1.2.3. *Any ball in a metric space M is open in M.*

Proof. Let $x \in B(a,\epsilon)$, then it can be checked that $B(x,\epsilon - d(a,x)) \subseteq B(a,\epsilon)$. □

Proposition 1.2.4 (Structure of Open Sets). *Any open set in a metric space M is a union of balls.*

Proof. Let U be open in M, let $x \in U$, then there must be δ_x such that $\{x\} \subseteq B(x,\delta_x) \subseteq U$. Hence taking union for all $x \in U$, we deduce that

$$\bigcup_{x\in U}\{x\} \subseteq \bigcup_{x\in U} B(x,\delta_x) \subseteq U \implies U = \bigcup_{x\in U} B(x,\delta_x). \qquad \square$$

Proposition 1.2.5. *The open subsets of a metric space M satisfy the following:*

(i) $\emptyset, X$ are open.

(ii) Arbitrary union of open subs is open.

(iii) Finite intersection of open sets are open.

Proof. (i) The reason for $\emptyset$ to be open is a kind of **vacuous truth** discussed in class, that is trivial. X is also open due to definition of ball, hence (i) is clear.

(ii) Let $\{U_\alpha\}_{\alpha\in A}$ be a collection of open subsets of M, then $u \in \bigcup_\alpha U_\alpha \implies u \in U_\alpha, \exists\alpha \implies B(u,\delta) \subseteq U_\alpha \subseteq \bigcup_\alpha U_\alpha$, for some $\delta > 0$.

(iii) It is enough to prove that U_1, U_2 open $\implies U_1 \cap U_2$ open. Let $u \in U_1 \cap U_2$, then there are δ_i, $i = 1,2$ such that $B(u,\delta_i) \subseteq U_i$. It follows that if we choose $\delta = \min\{\delta_1,\delta_2\} \le \delta_1,\delta_2$, then $B(u,\delta) \subseteq B(u,\delta_i)$, meaning that $B(u,\delta) \subseteq U_1 \cap U_2$. $\square$

Example 1.2.6. Let $A \subseteq (M,d)$, let $\epsilon > 0$, then

$$A^\epsilon := \{x \in M : d(x,a) < \epsilon, \exists a \in A\}$$

is open. You will be asked to give reason in presentation. $\square$

1.3 Continuity

Definition 1.3.1. A map $f : (X,d_X) \to (Y,d_Y)$ is said to be continuous at a if for any $\epsilon > 0$, there is a $\delta > 0$ such that

$$d_X(x,a) < \delta \implies d_Y(f(x),f(a)) < \epsilon.$$

Example 1.3.2. Let $f : (X,d_{\text{dis}}) \to (Y,d_Y)$, then f is automatically continuous. This is because given $a \in X$, then for any $\epsilon > 0$,

$$d_{\text{dis}}(x,a) < 1 \implies x = a \implies d_Y(f(x),f(a)) = 0 < \epsilon. \qquad \square$$

Example 1.3.3. The evaluation map $E(f) = f(0) : (C[0,1],d_\infty) \to \mathbb{R}$ is continuous since $|E(f) - E(g)| = |f(0) - g(0)| \le d_\infty(f,g)$, thus given $g \in C[0,1]$, for any $\epsilon > 0$,

$$d_\infty(f,g) < \epsilon \implies |E(f) - E(g)| < \epsilon. \qquad \square$$

Recall that the implication

$$d_X(x,a) < \delta \implies d_Y(f(x),f(a)) < \epsilon$$

is equivalent to saying that

$$x \in B_X(a,\delta) \implies \big(f(x) \in B_Y(f(a),\epsilon) \iff x \in f^{-1}\big(B_Y(f(a),\epsilon)\big)\big),$$

hence continuity of $f : (X,d_X) \to (Y,d_Y)$ at a is the same as saying for any $\epsilon > 0$, there is $\delta > 0$ such that

$$f(B_X(a,\delta)) \subseteq B_Y(f(a),\epsilon).$$

This observation leads us to the following interesting consequence which turns out to be a definition of continuous maps between two "topological spaces" (spaces on which we have defined what we mean by "open").

Proposition 1.3.4. *Consider $f : (X,d_X) \to (Y,d_Y)$, then the following are equivalent.*

(i) The map f is continuous.

(ii) $f^{-1}(U)$ is open in X for any open set $U \subseteq Y$.

Proof. Assume f is continuous, let U be open in Y, then $a \in f^{-1}(U) \implies f(a) \in U \implies \exists\epsilon > 0, B_Y(f(a),\epsilon) \subseteq U$. But from discussion above, there is a $\delta > 0$ such that $f(B_X(a,\delta)) \subseteq U \iff B_X(a,\delta) \subseteq f^{-1}(U)$.

Conversely, let $x \in X$ and $\epsilon > 0$ be given, by Proposition 1.2.3, $B_Y(f(x),\epsilon)$ is open, hence $V = f^{-1}(B_Y(f(x),\epsilon))$ is open. Clearly $x \in V$, and hence there is $\delta > 0$, $B_X(x,\delta) \subseteq V$, we are done. □

Proposition 1.3.4 provides us with an elegant way to prove the following that is usually proved in mathematical analysis course.

Corollary 1.3.5. *Composition of two continuous maps is continuous.*

Proof. Let $f : X \to Y$, $g : Y \to Z$ be continuous, then $g \circ f : X \to Z$ is continuous because for any U that is open in Z, $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U))$ is open in X. □

Example 1.3.6. The subset A of $\mathbb{R}^3$ defined by

$$A = \{(x,y,z) \in \mathbb{R}^3 : 1 < x + y^2 - z^3 + 10\sin x + 4\cos xy < 10\}$$

is open in $\mathbb{R}^3$ because $f(x,y,z) = x + y^2 - z^3 + 10\sin x + 4\cos xy$ is continuous and $A = f^{-1}\big((1,10)\big)$. □

Definition 1.3.7. Let M be a metric space, $\{x_n\}_{n=1}^{\infty}$ a sequence in M and $x \in M$. We say that $\{x_n\}_{n=1}^{\infty}$ converges to x (denoted by $\lim_{n\to\infty} x_n = x$) provided for any $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that

$$n > N \implies d(x_n,x) < \epsilon.$$

Proposition 1.3.8 (Sequential Continuity Theorem). *$f : (M_1,d_1) \to (M_2,d_2)$ is continuous at x_0 $\iff$ for every $\{x_n\}$ in M_1 that converges to x_0, $\lim_{n\to\infty} f(x_n) = f(x_0)$.*

Proof. The proof is essentially the same as real line case. $\square$

Hence once the sequence $\{x_n\}$ converges and f is continuous, the operation $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$ is valid as in the $\mathbb{R}$ case.

1.4 Limit Points

Definition 1.4.1. An **open neighborhood** of a point x in (M,d) is an open set U such that $x \in U$. A **deleted neighborhood** of x is an open neighborhood of x without x (that is, $U \setminus \{x\}$ but $x \in U$).

By *Proposition* 1.2.4, any open set must be a union of balls, hence it is convenient to consider the neighborhood with "minimal" size. Let's define

$$B'(a,\epsilon) = B(a,\epsilon) \setminus \{a\}.$$

Definition 1.4.2. Let M be a metric space. A point $x \in M$ is a **limit point** of $A \subseteq M$ if for all $\epsilon > 0$,

$$B'(x,\epsilon) \cap A \neq \emptyset.$$

Example 1.4.3. Consider $S := \{0\} \cup \{1 + \frac{1}{n} : n \in \mathbb{N}\}$. 0 is not a limit point because

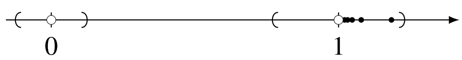


Figure 1.1: Example of limit points.

$B'(X,\epsilon) \cap A = \emptyset$ when $\epsilon < 1$. $1 \in S'$ for obvious reason, and there can't be any more (for the same reason as 0), hence $S' = \{1\}$. $\square$

Definition 1.4.4. Let $A \subseteq X$, the **derived set** of A is defined by

$$A' = \{x \in X : x \text{ is a limit point of } A\}.$$

Sometimes the derived set is also called the collection of accumulation/limit points.

Proposition 1.4.5. *Let A be a subset of metric space X, then*

$$x \in A' \iff \text{there are distinct } x_1, x_2, \cdots \in A \text{ such that } \lim_{n\to\infty} x_n = x.$$

Proof. Assume $x \in A'$, then there is $x_1 \in A$ such that $x_1 \in B'(x,1)$. Define $\{x_n\}$ inductively satisfying

$$x_n \in B'\left(x, \min\left\{d(x,x_{n-1}), \frac{1}{n}\right\}\right) \cap A$$

for $n \geq 2$, then clearly $x_1, x_2, \ldots$ are distinct with $\lim_{n\to\infty} x_n = x$.

The converse is obvious. $\square$

1.5 Closed Subsets

In mathematical analysis we have learnt that a closed set in $\mathbb{R}$ contains all its limit points, and a set is closed in $\mathbb{R}$ if and only if its complement is open in $\mathbb{R}$. They still hold in any metric space and it seems natural to define closed set as follows:

Definition 1.5.1. A subset of a metric space is **closed** if it contains all its limit point(s) (in other words, $A \supseteq A'$).

Example 1.5.2. (i) Singleton $\{a\}$ is closed since $\{a\}' = \emptyset \subseteq \{a\}$.

(ii) $[0,1]$ is closed but $[0,1)$ is not.

(iii) $\bigcup_{i=1}^{n}[0,1-\frac{1}{i}]$ is closed but $\bigcup_{i=1}^{\infty}[0,1-\frac{1}{i}]$ is not.

(iv) Both $\mathbb{Q}$ and $\mathbb{R}\setminus\mathbb{Q}$ are not closed.

(v) For any subset A of a metric space M, A' is closed. You will be asked to give reason in presentation.[1] $\square$

The following gives the relation between open sets and closed sets, it is another way to define closedness of a set in M.

Proposition 1.5.3. *A subset C of a metric space M is closed $\iff$ $M \setminus C$ is open.*

Proof. Assume C is closed, then $C \supseteq C'$. Take $x \in M \setminus C$, then in particular $x \notin C'$. By negating the definition of x being a limit point C, there is $\epsilon > 0$ such that $B'(x,\epsilon) \cap C = \emptyset$. Recall that $A \cap B = \emptyset \iff A \subseteq B^c$, hence

$$B'(x,\epsilon) \subseteq M \setminus C,$$

thus $x \in M \setminus C \implies B(x,\epsilon) \subseteq M \setminus C$.

Conversely, assume $M \setminus C$ is open. Let $x \in C'$, we claim that $x \in C$. For otherwise if $x \in M \setminus C$, then there is $\delta > 0$ such that $B(x,\delta) \subseteq M \setminus C$, definition of limit point tells us there is $x' \in B'(x,\delta) \cap C$, a contradiction. $\square$

Corollary 1.5.4. *A function $f : X \to Y$ between two metric spaces is continuous $\iff$ $f^{-1}(L)$ is closed in X whenever L is closed in Y.*

Proof. Recall that if $A \supseteq B$, $f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B)$ and $f^{-1}(Y) = X$. $\square$

Corollary 1.5.5. *The closed subsets of a metric space M satisfy the following.*

(i) $\emptyset, X$ are closed.

(ii) Arbitrary intersection of closed sets is closed.

(iii) Finite union of closed set is closed.

---

(1) This fails to be true in general topological space!

Proof. Recall Proposition 1.2.5 and De Morgan's laws:

$$M \setminus \bigcup_\alpha U_\alpha = \bigcap_\alpha (M \setminus U_\alpha) \quad \text{and} \quad M \setminus \bigcap_\alpha U_\alpha = \bigcup_\alpha (M \setminus U_\alpha). \qquad \square$$

Example 1.5.6. The n-sphere $S^n = \{(x_1,x_2,\ldots,x_{n+1}) \in \mathbb{R}^{n+1} : x_1^2 + x_2^2 + \cdots + x_{n+1}^2 = 1\}$ is closed since

$$f(x_1,x_2,\ldots,x_{n+1}) = x_1^2 + x_2^2 + \cdots + x_{n+1}^2$$

is continuous. $\square$

Finally we have a collection of symbols and definitions that are commonly used in point-set topology.

Definition 1.5.7. Let M be a metric space.

(i) x_0 is an **interior point** of S if there is $r > 0$, $B(x_0,r) \subseteq S$.

(ii) $S^\circ = \{x \in M : x \text{ is an interior point of } S\}$ is called an **interior** of S.

(iii) S' as previously defined is called **derived set**.

(iv) S is **open** in M if $S^\circ = S$.

(v) $\overline{S} = S \cup S'$ is called the **closure** of S in M.

(vi) S is **closed** if $\overline{S} = S$.

(vii) S is **dense** in M if $\overline{S} = M$.

(viii) A point $x \in M$ is an **boundary point** of S if for all $r > 0$, $B(x,r)$ has nonempty intersection with both S and $M \setminus S$.

(ix) $\partial S = \{x \in M : x \text{ is a boundary point of } S\}$.

For any subset S of a metric space, we expect after we fill in all the limit points of S to form a new set $\overline{S} := S \cup S'$, it becomes closed, and it can be shown easily by proposition Proposition 1.5.9. i.e., $\overline{S}$ is always a closed set containing S. Later on in exercises we will see that $\overline{S}$ is the smallest closed set containing S!

If S is closed, then $\overline{S} = S$. Conversely, if $\overline{S} = S$, then S is closed by the discussion above, hence

$$S \textit{ is closed} \iff \overline{S} = S.$$

Since $\overline{S} \supseteq S$ is always true, to show a set is closed it suffices to show $\overline{S} \subseteq S$, it is convenient to have an equivalence of "$x \in \overline{S}$".

Proposition 1.5.8 (Sequential Closure Theorem). *Let S be a subset of a metric space M, then*

$$x \in \overline{S} \iff \textit{there are } x_1, x_2, \cdots \in S,\ \lim_{n\to\infty} x_n = x.$$

Proof. Let $x \in \overline{S} := S \cup S'$. If $x \in S$, take $x_n = x$ for all $n \in \mathbb{N}$. If $x \in S'$, there is such a sequence by proposition Proposition 1.4.5.

Assume there are $x_1, x_2, \cdots \in S$ such that $\lim_{n\to\infty} x_n = x$. If $x_n = x$ for some $n \in \mathbb{N}$, then $x \in S$. Otherwise if $x_n \neq x$ for all $n \in \mathbb{N}$, we take $x_{n_1} = x_1$ and extract a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ which satisfies $d(x, x_{n_{k+1}}) < \min\{d(x, x_{n_k}), \frac{1}{k}\}$ for $k \geq 1$, then $x \in S'$. We conclude $x \in S \cup S' := \overline{S}$. □

Proposition Proposition 1.5.8 is another characterization of the closure $\overline{S}$ of S, it is sometimes useful when dealing with problems which involve continuous function. While the following characterization of closure can help us in many set theoretical problems.

Proposition 1.5.9. *Let S be a subset of a metric space M, then*

$$x \in \overline{S} \iff \textit{for any } \epsilon > 0,\ B(x, \epsilon) \cap S \neq \emptyset.$$

From this the closedness of $\overline{S}$ immediately follows. For a nice subset in M we can visualize the statement as in figure caption.3.

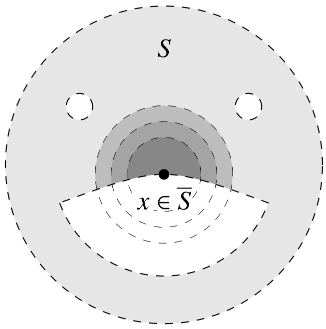


Figure 1.2: Points in the closure.

Proof. The ($\Rightarrow$) direction is clear by the definition of S'. For the ($\Leftarrow$) direction, we just need to consider two cases, namely, $x \in S$ (then we are done) and $x \notin S$ (pick $x_n \in B(x, \frac{1}{n}) \cap S$). □

Example 1.5.10. Let $\{S_\alpha\}_{\alpha\in A}$ be a collection of subsets of a metric space M, then

$$\overline{\bigcap_{\alpha\in A} S_\alpha} \subseteq \bigcap_{\alpha\in A} \overline{S_\alpha}.$$

This is because

$$\begin{aligned}
x \in \overline{\bigcap_{\alpha\in A} S_\alpha} &\iff \forall \epsilon > 0, B(x,\epsilon) \cap \left(\bigcap_{\alpha\in A} S_\alpha\right) \neq \emptyset \\
&\iff \forall \epsilon > 0, \bigcap_{\alpha\in A} (B(x,\epsilon) \cap S_\alpha) \neq \emptyset \\
&\implies \forall \epsilon > 0, B(x,\epsilon) \cap S_\alpha \neq \emptyset, \forall \alpha \in A \\
&\iff x \in \overline{S_\alpha}, \forall \alpha \in A
\end{aligned}$$

$$\iff x \in \bigcap_{\alpha \in A} \overline{S_\alpha}.$$

The equality can not hold in general. To see this, let $A = \mathbb{N}$ and for each $n \in \mathbb{N}$, define $S_n = \{\frac{1}{n}, \frac{1}{n+1}, \ldots\}$. Clearly $\bigcap_{n\in\mathbb{N}} S_n = \emptyset$ but $\bigcap_{n\in\mathbb{N}} \overline{S_n} = \{0\}$. $\square$

1.6 Exercises and Problems

Exercises

1.1. Let $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ be two normed vector spaces. Let $L(X,Y)$ denotes the collection of all continuous linear transformations from X to Y. Show that $\|T\| := \sup\{\|Tx\|_Y : x \in X, \|x\|_X = 1\}$ is a norm on $L(X,Y)$.

1.2. Let A be a subset of a metric space M, let $\epsilon > 0$, show that $A^\epsilon := \{x \in M : d(x,a) < \epsilon, \exists a \in A\}$ is open.

1.3. Prove the following properties of limit points.

(i) $A \subseteq B \implies A' \subseteq B'$.

(ii) $(A \cup B)' = A' \cup B'$ (hence $\overline{A \cup B} = \overline{A} \cup \overline{B}$).

(iii) $(A \cap B)' \subseteq A' \cap B'$ and the two sides may not be equal.

(iv) $A'' \subseteq A'$, and the two sides may not be equal.

Moreover, show that $(X \setminus A)'$ may not be equal to $X \setminus A'$. Also show that $(\bigcup_{i=1}^\infty A_i)' = \bigcup_{i=1}^\infty A_i'$ may not hold (therefore we don't expect $\overline{\bigcup_\alpha A_\alpha} = \bigcup_\alpha \overline{A_\alpha}$).

1.4. Let A be a subset of a metric space M. Define for $x \in M$,

$$d(x,A) = \inf\{d(x,a) : a \in A\}.$$

(a) Check that $d(x,A)$ is a continuous function in x.

(b) Prove that $d(x,A) = 0$ if and only if $x \in \overline{A}$.

(c) Show that any closed set in M is an intersection of a countable number of open set in M (called G_δ set).

1.5. Let M be a metric space and $A \subseteq M$.

(a) Prove that the interior A° of A is open in M and both the set A' and $\overline{A}$ are closed in M.

(b) Prove that A° is the union of all open sets in M contained in A. Prove that $\overline{A}$ is the intersection of all closed sets in M containing A.

Remark. This means A° is the largest open set in M contained in A and $\overline{A}$ is the smallest closed set in M containing A.

1.6. Let X be a metric space. A collection of subsets $\{A_i\}_{i\in I}$ is **locally finite** if for each $x \in X$, there is $\epsilon > 0$ such that $A_i \cap B(x,\epsilon) = \emptyset$ for all but finitely many A_i. Prove that the union of locally finite collection of **CLOSED** subsets is closed.

1.7. Let M_1, M_2 be metric spaces. Prove that the following are equivalent:

(i) $f : M_1 \to M_2$ is continuous.

(ii) for every $A \subseteq M_1$, we have $f(\overline{A}) \subseteq \overline{f(A)}$.

(iii) for every $B \subseteq M_2$, we have $\overline{f^{-1}(B)} \subseteq f^{-1}(\overline{B})$.

Problems

Definition 1.6.1. Define

$$C_c(\mathbb{R}) = \{f \in C(\mathbb{R}) : \text{exists a compact set } K \text{ in } \mathbb{R},\ f|_{\mathbb{R}\setminus K} = 0\}.$$

Examples are drawn in figure caption.6 (need not to share the same compact set). Such functions are said to have compact support. We then define $C_c^+(\mathbb{R}) = \{f \in C_c(\mathbb{R}) : f \geq 0\}$.

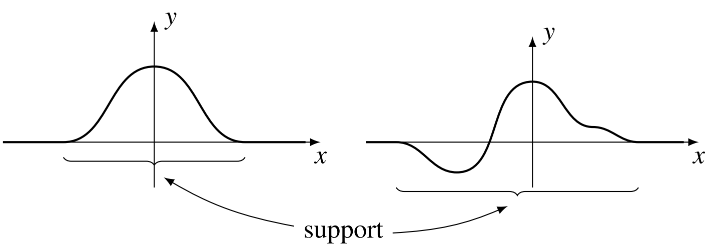


Figure 1.3: Compactly supported functions.

1.8. Given $f, g \in C_c^+(\mathbb{R})$, $g \not\equiv 0$, show that there are $a_i > 0$ and $s_j \in \mathbb{R}$ such that

$$f(x) \leq \sum_{j=1}^{n} a_j g(x - s_j), \quad \forall x \in \mathbb{R}.$$

Definition 1.6.2. Let X be a metric space and Λ a set of real numbers. A collection of open subsets of X $\{O_\lambda\}_{\lambda\in\Lambda}$ is said to be **normally ascending** provided for any $\lambda_1, \lambda_2 \in \Lambda$,

$$\overline{O_{\lambda_1}} \subseteq O_{\lambda_2} \text{ when } \lambda_1 < \lambda_2.$$

1.9. Let Λ be a dense subset of (a,b), where $a, b \in \mathbb{R}$, and $\{O_\lambda\}_{\lambda\in\Lambda}$ a normally ascending collection of open subsets of a metric space X. Define the function $f : X \to \mathbb{R}$ by setting $f = b$ on $X \setminus \bigcup_{\lambda\in\Lambda} O_\lambda$ and otherwise setting

$$f(x) = \inf\{\lambda \in \Lambda : x \in O_\lambda\}.$$

Show that $f : X \to [a,b]$ is continuous.

[**Hint:** $f : X \to [a,b]$ is continuous $\iff$ for each $c \in (a,b)$, the sets $\{x \in X : f(x) < c\}$ and $\{x \in X : f(x) > c\}$ are open. This result follows from the notion of subbase of metric topology on $(\mathbb{R}, |\cdot|)$.]

Definition 1.6.3. Let f be a real (or extended-real) valued function on a metric space X. If

$$\{x \in X : f(x) > \alpha\}$$

is open for every real α, f is said to be **lower semicontinuous**.

Remark. When X is any topological space, the notion of lower semicontinuity is defined in the same way. The simple example for such a function is the characteristic function of a open set in X (see definition Definition 3.3.2).

1.10. Suppose that X is a metric space, with metric d, and that $f : X \to [0,\infty]$ is lower semicontinuous, $f(p) < \infty$ for at least one $p \in X$. For $n = 1,2,3,\ldots$ and $x \in X$, define

$$g_n(x) = \inf\{f(p) + nd(x,p) : p \in X\}.$$

Prove that:

(i) $|g_n(x) - g_n(y)| \leq nd(x,y)$;

(ii) $0 \leq g_1 \leq g_2 \leq \cdots \leq f$ and

(iii) $\lim_{n\to\infty} g_n(x) = f(x)$, for all $x \in X$.

Chapter 2

Lebesgue Measure on $\mathbb{R}$

Throughout this text i,j,k and n are usually integers, when it is understood in the content we will simplify $*_{i=1}^{n}, *_{i=1}^{\infty}$ to $*_i / *$, where $* = \bigcup, \bigsqcup$ or $\sum$. We use simplified symbols to mean the union/series can be both finite or infinite.

2.1 Length of Open Sets

We have discussed what is meant by "open", we now show that open sets are disjoint unions (denoted by $\sqcup$) of countably many open intervals, and the decompositions are unique.

Proposition 2.1.1. *Any open subsets of $\mathbb{R}$ is a union of countably many pairwise disjoint open intervals. Moreover, the decomposition is unique.*

Proof. We first explain the uniqueness of the decomposition. Suppose an open set on $\mathbb{R}$ has two decompositions $\bigsqcup U_i = \bigsqcup V_i$, where U_i, V_j are intervals, then $U_i = \bigsqcup_j (U_i \cap V_j)$. But there is one and only one $U_i \cap V_j$ can be nonempty in the union (otherwise U_i is split into at least two intervals), and hence $U_i = U_i \cap V_j$, for some j. But for the same reason, $V_j = V_j \cap U_i$, hence $U_i = V_j$.

By Proposition 1.2.4 any open set O is a union of balls, i.e., we can write $O = \bigcup_{x \in O} (a_x, b_x)$, where $x \in (a_x, b_x)$. We now extend our intervals as large as possible. Since $L_x := \{a \in \mathbb{R} : (a,x] \subseteq O\}$ and $R_x := \{b \in \mathbb{R} : [x,b) \subseteq O\}$ are nonempty, define

$$a'_x = \inf L_x \quad \text{and} \quad b'_x = \sup R_x$$

(can be $\mp\infty$) and write $I_x = (a'_x, b'_x)$. We show that $I_x \subseteq O$. Pick a $y \in I_x$ and let's first assume $y < x$. Then as $a'_x = \inf L_x < y$, there is small enough a in L_x such that $a < y$ and $(a,x] \subseteq O$, hence $y \in O$. The case that $y > x$ is essentially the same. Now we can write $O = \bigcup_{x \in O} I_x$.

We show that distinct intervals in the union are disjoint. Assume there are $x, y \in O$, $x < y$, such that $I_x \neq I_y$. If $I_x \cap I_y \neq \emptyset$, then $I_x \cup I_y$ is an open interval. As $I_x \neq I_y$, $a'_x \neq a'_y$ or $b'_x \neq b'_y$, either one of them is a contradiction[(1)]. We conclude $I_x \cap I_y = \emptyset$.

[(1)] Let's say if $a'_x < a'_y$, then since $I_x \cup I_y \subseteq O$ is an interval, $a_x \in L_y$, but $a'_y := \inf L_y \leq a'_x$, a contradiction.

Let $\{I_\alpha : \alpha \in A\}$ be the collection of all distinct elements in $\{I_x : x \in O\}$. For each interval I_α we choose a value $r_\alpha \in I_\alpha \cap \mathbb{Q}$ and construct a map $I_\alpha \mapsto r_\alpha$, which is injective, hence $\{I_\alpha : \alpha \in A\}$ is countable. □

Due to Proposition 2.1.1 we can now define the "length" of any open set as follows:

Definition 2.1.2. The **length** of an open set $U = \bigsqcup(a_i, b_i)$ is

$$\lambda(U) = \sum(b_i - a_i).$$

If an open set contains an unbounded open interval, its length is ∞.

As in probability, we expect the length of open sets should satisfy

$$\lambda(U \cup V) = \lambda(U) + \lambda(V) - \lambda(U \cap V), \tag{2.1.3}$$

which is indeed true. To justify the equality, we notice that the proof is complicated when the open sets are unions of countably infinitely many intervals. Hence we divide the proof into two cases, finite union and infinite union. The finite version is treated in the following proposition.

Proposition 2.1.4. *Let U,V be finite union of open finite intervals, then*

$$\lambda(U \cup V) = \lambda(U) + \lambda(V) - \lambda(U \cap V).$$

Proof. Let $E = \{a_0, \dots, a_N\}$, $a_1 < \cdots < a_N$ be the collection of distinct end points of intervals contained in A and B. Let $I_i = (a_i, a_{i-1})$, $i = 1, 2, \dots, N$. Then $U, V, U \cup V, U \cap V$ are union of intervals in $\mathcal{I} := \{I_i\}_{i=1}^N$ except possibly finitely many points which are common end points of two adjacent intervals in $\mathcal{I}$. Hence

$$\begin{aligned}
\lambda(U \cap V) &= \sum_{I \in \mathcal{I}, I \subseteq U \cup V} \lambda(I) && (2.1.5)\\
&= \sum_{I \in \mathcal{I}, I \subseteq U} \lambda(I) + \sum_{I \in \mathcal{I}, I \subseteq V} \lambda(I) - \sum_{I \in \mathcal{I}, I \subseteq U \cap V} \lambda(I) && (2.1.6)\\
&= \lambda(U) + \lambda(V) - \lambda(U \cap V). && (2.1.7)
\end{aligned}$$

(2.1.5) and (2.1.7) are due to definition of length of open sets and the fact that $\lambda(I_j \sqcup I_{j+1}) = a_j - a_{j-1} + a_{j+1} - a_j = a_{j+1} - a_{j-1} = \lambda(a_{j-1}, a_{j+2})$. (2.1.6) is due to the fact that the length of I contained in $U \cap V$ is double counted in $\sum_{I \in \mathcal{I}, I \subseteq U} \lambda(I) + \sum_{I \in \mathcal{I}, I \subseteq V} \lambda(I)$. □

We now show that λ is "countably monotone" ((i) of Proposition 2.1.8) and λ satisfies equation (2.1.3) in general.

Proposition 2.1.8. *The length of open sets has the following properties.*

(i) If $U \subseteq \bigcup V_i$, then $\lambda(U) \leq \sum \lambda(V_i)$.

(ii) $\lambda(U \cup V) = \lambda(U) + \lambda(V) - \lambda(U \cap V)$.

Proof. (i) If there is $\lambda(V_i) = \infty$, done. Assume for all i, $\lambda(V_i) < \infty$. Write $U = \bigsqcup(a_i, b_i)$, since V_i's are open, write V_i as a disjoint union of open intervals, collect all such intervals (c_i, d_i) and express $\bigcup V_i = \bigcup (c_i, d_i)$. Fix an $n \in \mathbb{N}$ and choose $\epsilon > 0$ small, define

$$K_n = \bigsqcup_{i=1}^{n} [a_i + \epsilon, b_i - \epsilon].$$

Clearly K_n is compact and $K_n \subseteq U \subseteq \bigsqcup (c_i, d_i)$, hence there is a $k_n \in \mathbb{N}$ such that

$$K_n \subseteq \bigcup_{i=1}^{k_n} (c_i, d_i).$$

For union of finitely many intervals, it is easy to see that

$$\sum_{i=1}^{n} (b_i - a_i) - 2n\epsilon = \sum_{i=1}^{n} [(b_i - \epsilon) - (a_i + \epsilon)] \le \sum_{i=1}^{k_n} (d_j - c_j) \le \sum (d_j - c_j),$$

letting $\epsilon \to 0^+$, we infer from the last inequality that $\sum_{i=1}^{n} (b_i - a_i) \le \sum (d_j - c_j)$, but this is true for each n, hence $\lambda(U) \le \sum (d_j - c_j) = \sum \lambda(V_i)$.

(ii) If $\lambda(U)$ or $\lambda(V)$ is unbounded, then we are done. Assume now $\lambda(U), \lambda(V) < \infty$, let $U = \bigsqcup (a_i, b_i), V = \bigsqcup (c_i, d_i)$, then construct the finite unions

$$U_n = \bigsqcup_{i=1}^{n} (a_i, b_i) \quad \text{and} \quad V_n = \bigsqcup_{i=1}^{n} (c_i, d_i).$$

For finite union of intervals it is proved Proposition 2.1.4 that

$$\lambda(U_n \cup V_n) = \lambda(U_n) + \lambda(V_n) - \lambda(U_n \cap V_n). \tag{2.1.9}$$

Since both $\lambda(U)$ and $\lambda(V)$ are bounded, given $\epsilon > 0$, there is an N such that when $n > N$,

$$0 \le \lambda(U) - \lambda(U_n) = \lambda(U \setminus U_n) = \sum_{i>n} \lambda((a_i, b_i)) < \epsilon,$$

$$0 \le \lambda(U) - \lambda(V_n) = \lambda(V \setminus V_n) = \sum_{i>n} \lambda((c_i, d_i)) < \epsilon.$$

It is left as exercise to show that

$$U \cup V \subseteq (U_n \cup V_n) \cup (U \setminus U_n) \cup (V \setminus V_n),$$

$$U \cap V \subseteq (U_n \cap V_n) \cup (U \setminus U_n) \cup (V \setminus V_n),$$

it follows from (i) of Proposition 2.1.8 that when $n > N$,

$$\lambda(U \cup V) - \lambda(U_n \cup V_n) < 2\epsilon \quad \text{and} \quad \lambda(U \cap V) - \lambda(U_n \cap V_n) < 2\epsilon.$$

These prove $\lim_{n\to\infty} \lambda(U_n \cup V_n) = \lambda(U \cup V)$ and $\lim_{n\to\infty} \lambda(U_n \cap V_n) = \lambda(U \cap V)$, hence we get desired result from equation (2.1.9). □

2.2 Length of Closed Sets

For a compact set K and any bounded open set U containing K, we have a decomposition $U = K \sqcup (U \setminus K)$, and thus $U \setminus (U \setminus K) = K$. Note that both U and $U \setminus K$ are open for which we have defined their length, we have the following natural definition.

Definition 2.2.1. The **length** of a compact subset $K \subseteq \mathbb{R}$ is

$$\lambda(K) = \lambda(U) - \lambda(U \setminus K),$$

where U is a bounded open set containing K.

$\lambda(K)$ is well-defined (that is, $\lambda(K)$ is independent of the choice of bounded open set $U \supseteq K$). To see this, let U, V be two bounded open sets containing K, then clearly $U \supseteq U \cap V \supseteq K$, hence by Proposition 2.1.8,

$$\begin{aligned}\lambda(U) &= \lambda((U \setminus K) \cup (U \cap V)) \\ &= \lambda(U \setminus K) + \lambda(U \cap V) - \lambda((U \setminus K) \cap (U \cap V)) \\ &= \lambda(U \setminus K) + \lambda(U \cap V) - \lambda((U \cap V) \setminus K),\end{aligned}$$

hence

$$\lambda(U) - \lambda(U \setminus K) = \lambda(U \cap V) - \lambda((U \cap V) \setminus K).$$

Interchanging U and V, we immediately we find that $\lambda(U) - \lambda(U \setminus K) = \lambda(V) - \lambda(V \setminus K)$, so the definition is not ambiguous.

Another weird proof for well-definedness can be found in Kin Li's MATH301 notes, page 105. For completeness, we also define the length of any closed sets.

Definition 2.2.2. The length of a closed set L is

$$\lambda(L) = \lim_{x \to +\infty} \lambda(L \cap [-x, x]).$$

2.3 Lebesgue Measure

We try to extend the length to sets other than open and closed ones by approximating any set from outside by open sets and from inside by compact subsets. These approximations give upper and lower bounds of the "length" of the set. When two bounds are equal, there is no ambiguity on the length of the set, so that the length is well-defined.

The following is the definition of Lebesgue measurability and Lebesgue measure, however, after Theorem 2.4.5 and Proposition 2.7.4, we may replace it by Definition 2.5.1 and Definition 2.7.5.

Definition 2.3.1.

(i) The **Lebesgue outer measure** of a subset $A \subseteq \mathbb{R}$ is

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

(ii) The **Lebesgue inner measure** of a subset $A \subseteq \mathbb{R}$ is

$$m_*(A) = \sup\{\lambda(K) : K \subseteq A,\ K \text{ compact}\}.$$

(iii) A bounded set A is said to be **Lebesgue measurable** if $m_*(A) = m^*(A)$ and the common value is the **Lebesgue measure**, denoted by $m(A)$.

(iv) An unbounded set A is said to be **Lebesgue measurable** if $A \cap [a, b]$ is measurable for every $a \leq b$. In this case the **Lebesgue measure** of A is

$$m(A) = \lim_{x \to +\infty} m(A \cap [-x, x]).$$

Throughout this and next chapters “measurable” means “Lebesgue measurable”, for short.

Proposition 2.3.2. *The outer and inner measures have the following properties:*

(i) $0 \le m_*(A) \le m^*(A)$.

(ii) $A \subseteq B \implies m_*(A) \le m_*(B)$ *and* $m^*(A) \le m^*(B)$.

(iii) $m^*(\bigcup A_i) \le \sum m^*(A_i)$.

Proof. (i) If $m^*(A) = \infty$, we are done. If $m^*(A) < \infty$, we can choose open $U \supseteq A$ that has finite length, and the inequality follows from $\lambda(K) = \lambda(U) - \lambda(U - K) \le \lambda(U)$ for any compact subset $K \subseteq A$.

(ii) To get the first inequality, let $L \subseteq A$ and $K \subseteq B$ be compact, then $L \cup K \subseteq A \cup B = B$, hence

$$\lambda(L) \overset{\text{(why?)}}{\le} \lambda(L \cup K) \le m_*(B),$$

taking the supremum over all compact $L \subseteq A$, we get $m_*(A) \le m_*(B)$. The second one can be similarly proved.

(iii) If $m^*(A_i) = \infty$ for some i, we are done. Assume for all i, $m^*(A_i) < \infty$, let $\epsilon > 0$, for each A_i there is an open set $U_i \supseteq A_i$ such that $\lambda(U_i) - m^*(A_i) < \frac{\epsilon}{2^i}$, then clearly $\bigcup U_i \supseteq \bigcup A_i$. By letting $U = \bigcup V_i$ and setting $V_i = A_i$ in (i) of Proposition 2.1.8, we get

$$m^*\Big(\bigcup A_i\Big) \le \lambda\Big(\bigcup U_i\Big) \le \sum \lambda(U_i) \le \sum \Big(m^*(A_i) + \frac{\epsilon}{2^i}\Big) \le \sum m^*(A_i) + \epsilon.$$

As $\epsilon > 0$ is arbitrary, we are done. □

The following is an immediate consequence.

Corollary 2.3.3. *If* $m^*(A) = 0$*, then* A *is measurable with* $m(A) = 0$*. Moreover, any subset of a set of measure zero is also a set of measure zero.*

Proposition 2.3.4.

(i) Intervals are measurable with the usual length as its measure.

(ii) Let U *be open, then* $\lambda(U) = m_*(U) = m^*(U)$.

Proof. (i) The measurability of any kind of bounded interval $A = \langle a, b\rangle$ can be obtained by

$$[a + \epsilon, b - \epsilon] \subseteq A \subseteq (a - \epsilon, b + \epsilon),$$

it gives us the estimate $(b - a) - 2\epsilon \le m_*(A) \le m^*(A) \le (b - a) + 2\epsilon$, for all $\epsilon > 0$. For unbounded interval I, $I \cap [-x, x]$ is measurable. In any case, let $I = \mathbb{R}$, $(-\infty, a\rangle$ or $\langle a, +\infty)$, we get

$$m(I) = \lim_{x \to +\infty} m(I \cap [-x, x]) = +\infty.$$

(ii) Let's assume U is a disjoint union of bounded open intervals first. Let $\epsilon > 0$ and $U = \bigsqcup (a_i, b_i)$, consider

$$\bigsqcup_{i=1}^{n} [a_i + \epsilon, b_i - \epsilon] \subseteq U \subseteq U,$$

which tells us $\sum_{i=1}^{n}(b_i - a_i) - 2n\epsilon \le m_*(U) \le m^*(U) \le \sum(b_i - a_i)$, and we get desired equality by letting $\epsilon \to 0^+$ and then taking $n = N$ (if $\bigsqcup = \bigsqcup_{i=1}^{N}$) or $n \to +\infty$ (if $\bigsqcup = \bigsqcup_{i=1}^{\infty}$). If U contains an unbounded interval, say $U = \bigsqcup(a_i, b_i) \sqcup (a, +\infty)$, then as above the two sides approximation (k large)

$$\bigsqcup_{i=1}^{n}[a_i + \epsilon, b_i - \epsilon] \sqcup [a + \epsilon, k] \subseteq U \subseteq U$$

implies $\sum_{i=1}^{n}(b_i - a_i) + (k - a) - (2n+1)\epsilon \le m_*(U) \le m^*(U) \le +\infty$, the result follows from first letting $\epsilon \to 0^+$ and then $k \to +\infty$. $\square$

The following shows that outer measure is **translation invariant**.

Proposition 2.3.5. *Let $E \subseteq \mathbb{R}$ and $x \in \mathbb{R}$, then define $E + x = \{e + x : e \in E\}$, we have*

$$m^*(E + x) = m^*(E).$$

Proof. Let $O \supseteq E$ be open, then clearly $O + x \supseteq E + x$ is also open, hence $m^*(E + x) \le m^*(O + x) = \lambda(O + x) = \lambda(O)$, taking infimum of $\lambda(O)$ over all open $O \supseteq E$, one has

$$m^*(E + x) \le m^*(E).$$

We repeat the process for the translation $-x$, obtaining the reverse inequality. $\square$

2.4 Carathéodory Theorem

Before we prove Carathéodory theorem we need two technical results. In mathematical analysis, given a sequence of real numbers $\{x_n\}$, denote

$$\mathcal{L} = \{\ell \in [-\infty, \infty] : \text{exists } x_{n_k},\ \lim_{k\to\infty} x_{n_k} = \ell\},$$

we denote $\underline{\lim}_{n\to\infty} x_n = \inf \mathcal{L}$ and $\overline{\lim}_{n\to\infty} x_n = \sup \mathcal{L}$, we have for two sequences of real numbers:

$$\underline{\lim}_{n\to\infty}(a_n + b_n) \le \underline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} b_n \le \overline{\lim}_{n\to\infty}(a_n + b_n).$$

The following lemma shares the same pattern.

Lemma 2.4.1. *If A and B are disjoint, then*

$$m_*(A \sqcup B) \le m_*(A) + m^*(B) \le m^*(A \sqcup B).$$

Proof. Consider the right inequality first. If $m^*(A \sqcup B) = \infty$, we have nothing to prove.

Assume now $m^*(A \sqcup B) < \infty$. As usual to get a relation with outer measure and inner measure we approximate $A \sqcup B$ from outside and approximate A from inside. Let $\epsilon > 0$, then we can find an open $U \supseteq A \sqcup B$ and a compact $K \subseteq A$ such that

$$\lambda(U) - m^*(A \sqcup B) < \epsilon \quad \text{and} \quad m_*(A) - \lambda(K) < \epsilon.$$

Recall that $\lambda(K) = \lambda(U) - \lambda(U \setminus K)$, and $U \setminus K \supseteq B$ is open, hence adding them up and rearranging terms, we deduce that

$$m^*(A \sqcup B) + 2\epsilon > m_*(A) + \lambda(U) - \lambda(K) = m_*(A) + \lambda(U \setminus K) \geq m_*(A) + m^*(B).$$

We let $\epsilon \to 0^+$ to get desired inequality.

We now consider the left inequality, let there be a compact $L \subseteq A \sqcup B$ and an open $V \supseteq B$, then $L \setminus V$ is a compact subset contained in A, we claim that

$$\lambda(L) \leq \lambda(L \setminus V) + \lambda(V). \tag{2.4.2}$$

Let $O \supseteq L$ have finite length, the inequality (2.4.2) is the same as

$$\lambda(O) - \lambda(O \setminus L) \leq \lambda(O) - \lambda(O \setminus (L \setminus V)) + \lambda(V) \iff \lambda(O \setminus (L \setminus V)) \leq \lambda(V) + \lambda(O \setminus L)$$

but this is true since $O \setminus (L \setminus V) = (O \cap V) \cup (O \setminus L) \subseteq V \cup (O \setminus L)$.

From (2.4.2),

$$\lambda(L) \leq \lambda(L \setminus V) + \lambda(V) \leq m_*(A) + \lambda(V),$$

this is true for all open V containing B and compact L contained in $A \sqcup B$, taking supremum first and then infimum (or reverse the order), we get

$$m_*(A \sqcup B) \leq m_*(A) + m^*(B). \qquad \square$$

Lemma 2.4.3. *Let B be an unbounded measurable subset of $\mathbb{R}$ and U be open with finite length, then*

$$m^*(B \cap U) \leq m_*(B \cap U).$$

Proof. Write $U = \bigsqcup_i (a_i, b_i)$, then given $\epsilon > 0$, there is an N such that $\lambda(U) - \sum_{i=1}^N \lambda(a_i, b_i) < \epsilon$. Let's define $S = \bigsqcup_{i=1}^N (a_i, b_i)$, our goal is to find the length of inner approximation of $U \cap B$ as an upper bound of $m^*(U \cap B)$.

$$\begin{aligned}
m^*(U \cap B) &\leq m^*(S \cap B) + m^*((U \setminus S) \cap B) < m^*(S \cap B) + \epsilon \\
&\leq \sum_{i=1}^N m^*([a_i, b_i] \cap B) + \epsilon \\
&\leq \sum_{i=1}^N \left(m^* \left(\left[a_i + \frac{\epsilon}{2^{i+1}}, b_i - \frac{\epsilon}{2^{i+1}} \right] \cap B \right) + \frac{\epsilon}{2^i} \right) + \epsilon \\
&< \sum_{i=1}^N m^* \Bigg(\underbrace{\left[a_i + \frac{\epsilon}{2^{i+1}}, b_i - \frac{\epsilon}{2^{i+1}} \right] \cap B}_{:= L_i} \Bigg) + 2\epsilon. \\
&= \sum_{i=1}^N m_*(L_i) + 2\epsilon.
\end{aligned} \tag{2.4.4}$$

For each bounded subset $L_i \subseteq U \cap B$, we can find a compact $K_i \subseteq L_i$ such that $m_*(L_i) < \lambda(K_i) + \frac{\epsilon}{N}$, hence $\sum_{i=1}^N m_*(L_i) < \sum_{i=1}^N \lambda(K_i) + \epsilon \overset{\text{(why?)}}{=\!=\!=} \lambda(\bigsqcup_{i=1}^N K_i) + \epsilon$, by (2.4.4),

$$m^*(U \cap B) < \sum_{i=1}^N m_*(L_i) + 2\epsilon < \lambda \left(\bigsqcup_{i=1}^N K_i \right) + 3\epsilon \leq m_*(U \cap B) + 3\epsilon,$$

we complete the proof by letting $\epsilon \to 0^+$. $\square$

We are in a position to prove one of the main theorems of this section, which basically says that a set is measurable if and only if it and its complement can be used to "split" the outer measure of any set.

Theorem 2.4.5 (Carathéodory). *A set A is measurable if and only if the Carathéodory condition*

$$m^*(X) = m^*(X \cap A) + m^*(X \setminus A)$$

holds for any set X.

Proof. ($\Leftarrow$) Let X be a bounded interval $[a,b]$ with $a \le b$, then by 2.4,

$$\begin{aligned} \boxed{m^*([a,b] \cap A)} + m^*([a,b] \setminus A) &= m^*([a,b]) \\ &= m_*([a,b]) \\ &\le \boxed{m_*([a,b] \cap A)} + m^*([a,b] \setminus A), \end{aligned}$$

this implies $m^*(A \cap [a,b]) \le m_*(A \cap [a,b])$. Hence A is measurable, no matter it is bounded or not.

($\Rightarrow$) By (iii) of Proposition 2.3.2, it suffices to show that $m^*(X) \ge m^*(X \cap A) + m^*(X \setminus A)$ for any set X. We first observe that when $m^*(X) = \infty$, the equality is trivial. So we assume now $m^*(X) < \infty$.

We first prove any open set satisfies Carathéodory condition. Let O be open. Given $\epsilon > 0$ there is an open $U \supseteq X$ such that $\lambda(U) - m^*(X) < \epsilon$. Thus

$$\begin{aligned} m^*(X) + \epsilon > \lambda(U) = m^*(U) &\ge m_*(U \cap O) + m^*(U \setminus O) && (2.4.6) \\ &= m^*(U \cap O) + m^*(U \setminus O) && (2.4.7) \\ &\ge m^*(X \cap O) + m^*(X \setminus O), \end{aligned}$$

where (2.4.7) follows from $m_*(U \cap O) = m^*(U \cap O)$ by (ii) of Proposition 2.3.4. We then let $\epsilon \to 0^+$ to get: For any set X and open set O,

$$m^*(X) \ge m^*(X \cap O) + m^*(X \setminus O). \tag{2.4.8}$$

Now we try to prove (2.4.8) is also true when O is replaced by A. We first claim that

$$m_*(U \cap A) \ge m^*(U \cap A). \tag{2.4.9}$$

When A is unbounded, (2.4.9) is true immediately by Lemma 2.4.3. When A is bounded, we let $X = A$ and $O = U$ in (2.4.8), then by 2.4,

$$\boxed{m^*(A \cap U)} + m^*(A \setminus U) = m^*(A) = m_*(A) \le \boxed{m_*(A \cap U)} + m^*(A \setminus U),$$

we conclude (2.4.9) for any measurable A, and thus if we redo (2.4.6),

$$\begin{aligned} m^*(U) &\ge m_*(U \cap A) + m^*(U \setminus A) \\ &\ge m^*(U \cap A) + m^*(U \setminus A) \ge m^*(X \cap A) + m^*(X \setminus A). \end{aligned}$$

We complete the proof by letting $\epsilon \to 0^+$. $\square$

2.5 The sigma-Algebra of Lebesgue Measurable Sets

Because of Theorem 2.4.5 we see that measurability of subsets in $\mathbb{R}$ can be characterized by the outer measure alone. Let's copy and paste that theorem as our new definition of measurability.

Definition 2.5.1. A set E is said to be **measurable** provided for any set X,

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

In Definition 2.7.5 we will redefine **Lebesgue measure** of A, $m(A)$, to be the restriction of m^* to the collection of measurable subsets, but not now.

For convenience we may write E^c to denote the relative complement of E in $\mathbb{R}$. Note that $(E^c)^c = E$, hence by the definition of measurability, E is measurable if and only if E^c is measurable.

As pointed out in the ($\Rightarrow$) direction of the Theorem 2.4.5's proof, a set is measurable if and only if $m^*(X) \geq m^*(X \cap A) + m^*(X \setminus A)$ for any set X. It provides us with a handy criterion to check measurability of many conceivable sets, for example, countable union and intersection of measurable sets. We will mention it one by one.

Proposition 2.5.2. *The union of a finite collection of measurable sets is measurable.*

Proof. It suffices to prove that when A,B are measurable, so is $A \cup B$. Let T be any subset of $\mathbb{R}$, then by the set equalities

$$(T \cap A) \cup (T \cap A^c \cap B) = T \cap (A \cup B) \quad \text{and} \quad A^c \cap B^c = (A \cup B)^c,$$

one has

$$\begin{aligned}
m^*(T) &= m^*(T \cap A) + m^*(T \cap A^c) \\
&= m^*(T \cap A) + m^*(T \cap A^c \cap B) + m^*(T \cap A^c \cap B^c) \\
&\geq m^*((T \cap A) \cup (T \cap A^c \cap B)) + m^*(T \cap A^c \cap B^c) \\
&= m^*(T \cap (A \cup B)) + m^*(T \cap (A \cup B)^c).
\end{aligned}$$

□

Proposition 2.5.3. *Let A be any set and $\{E_k\}_{k=1}^n$ be a finite disjoint collection of measurable sets. Then*

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

In particular,

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

and we call m^* ***finitely additive****.*

Proof. We prove by induction on n, The case that $n = 1$ is clear. Assume it is true for $n-1$, then

$$m^*\left(A \cap \left(\bigcup_{k=1}^n E_k\right)\right) = m^*\left(A \cap \left(\bigcup_{k=1}^n E_k\right) \cap E_n\right) + m^*\left(A \cap \left(\bigcup_{k=1}^n E_k\right) \cap E_n^c\right)$$

$$= m^*\left(A \cap \left(\bigcup_{k=1}^{n} E_k \cap E_n\right)\right) + m^*\left(A \cap \left(\bigcup_{k=1}^{n-1} E_k\right)\right)$$

$$= m^*(A \cap E_n) + \sum_{k=1}^{n-1} m^*(A \cap E_k). \qquad \square$$

Proposition 2.5.4. *The countable union of measurable sets is measurable.*

Proof. Let $\{E_k\}_{k=1}^{\infty}$ be a collection of measurable sets. Define $F_1 = E_1$ and $F_k = E_k \setminus \bigcup_{i=1}^{k-1} E_i$ for $k \geq 2$, then $\{F_k\}_{k=1}^{\infty}$ is a disjoint collection of measurable sets and $\bigcup_{k=1}^{\infty} F_k = \bigcup_{k=1}^{\infty} E_k$. Let A be any set and define $E = \bigcup_{k=1}^{\infty} E_k$, then by Proposition 2.5.3,

$$m^*(A) = m^*\left(A \cap \left(\bigcup_{k=1}^{n} F_k\right)\right) + m^*\left(A \cap \left(\bigcup_{k=1}^{n} F_k\right)^c\right) \geq \sum_{k=1}^{n} m^*(A \cap F_k) + m^*(A \cap E^c)$$

for each $n \in \mathbb{N}$, hence from Proposition 2.3.2,

$$m^*(A) \geq \sum_{k=1}^{\infty} m^*(A \cap F_k) + m^*(A \cap E^c) \geq m^*(A \cap E) + m^*(A \cap E^c). \qquad \square$$

Corollary 2.5.5. *The countable intersection of measurable sets is measurable.*

Proof. It follows from De Morgan's laws. $\square$

A collection of subsets of $\mathbb{R}$ is called an **algebra** if it contains $\mathbb{R}$ and is closed under relative complement and finite union. The prefix σ refers to properties related to countable union. For example, a countable union of closed sets is called a F_σ set. A countable intersection of open sets is called a G_δ set. A countable union of G_δ sets is called a $G_{\delta\sigma}$ set.

An algebra is called a σ-algebra, as defined in Definition 2.10.1, if it is further closed under **countable union**. We summarize this section by noting that the collection of (Lebesgue) measurable subsets $\mathcal{L}$ is a σ-algebra.

2.6 Approximation of Lebesgue Measurable Sets

Measurable sets possess the **excision property**, that is, if $E \subseteq A$ is measurable and has finite outer measure, then

$$m^*(A \setminus E) = m^*(A) - m^*(E),$$

this follows from the definition of measurability of E. The validity of transposing the term $m^*(E)$ requires it be finite. Note that we have used the following fact many times:

If E has finite outer measure, for any $\epsilon > 0$ there is an open set O such that

$$m^*(O) - m^*(E) < \epsilon.$$

In fact the above can also be rewritten as $m^*(O \setminus E) < \epsilon$, and this expression also makes sense even if E has unbounded outer measure. This observation formulates another useful criterion of measurability.

Theorem 2.6.1. *Let $E \subseteq \mathbb{R}$, the following are equivalent:*

(i) E is measurable.

(Outer Approximation by Open Sets and G_δ Sets)

(ii) For each $\epsilon > 0$, there is an open O containing E for which $m^*(O \setminus E) < \epsilon$.

(iii) There is a G_δ set G containing E for which $m^*(G \setminus E) = 0$.

(Inner Approximation by Closed Sets and F_σ Sets)

(iv) For each $\epsilon > 0$, there is a closed F contained in E for which $m^*(E \setminus F) < \epsilon$.

(v) There is an F_σ set F contained in E for which $m^*(E \setminus F) = 0$.

It suffices to show that (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii), while (ii) $\Leftrightarrow$ (iv) and (iii) $\Leftrightarrow$ (v) easily follow from the observation that $A \setminus B = B^c \setminus A^c$.

Proof. (i) $\Rightarrow$ (ii) Assume E is measurable, if $m^*(E) < \infty$, then (i) follows from early discussion in this section.

If $m^*(E) = \infty$, let's say $E = \bigsqcup_n E_n$ with $m^*(E_n) < \infty$, then for any $\epsilon > 0$, there is an open O_n such that $m^*(O_n \setminus E_n) < \epsilon/2^n$. Define $O = \bigcup_n O_n$ and hence

$$m^*(O \setminus E) = m^*(\bigcup_n (O_n \setminus E)) \le m^*(\bigcup_n (O_n \setminus E_n)) \le \sum_n m^*(O_n \setminus E_n) < \epsilon.$$

(ii) $\Rightarrow$ (iii) Since for each n there is an open $O_n \supseteq E$ such that $m^*(O_n \setminus E) < \frac{1}{n}$. If we construct $G = \bigcap_{n=1}^\infty O_n$, then $m^*(G \setminus E) \le m^*(O_n \setminus E) < \frac{1}{n}$, for all $n \in \mathbb{N}$.

(iii) $\Rightarrow$ (i) Since there is a G_δ set $G \supseteq E$ such that $m^*(G \setminus E) = 0$, hence $G \setminus E$ is measurable, this implies $E = G \cap (G \setminus E)^c$ is measurable. ☐

Remark. In the proof of (i) $\Rightarrow$ (ii) the decomposition of E is not necessarily disjoint. For example, the decomposition $E = \bigcup_{n=1}^\infty (E \cap [-n, n])$ will also do.

Example 2.6.2. This is a simple application of G_δ set. Let $x \in \mathbb{R}$ and $A \subseteq \mathbb{R}$, define $A + x = \{a + x : a \in A\}$, it is natural to ask if A is measurable, is the translation $A + x$ also measurable? The answer is positive. Let there be a G_δ set $G \supseteq A$ such that $A = G \setminus (G \setminus A)$ with $m(G \setminus A) = 0$, then

$$A + x = G \setminus (G \setminus A) + x = (G + x) \setminus (G \setminus A + x).$$

However, by Proposition 2.3.5 outer measure is translation invariant, hence $m^*(G \setminus A + x) = m^*(G \setminus A) = 0$. But $G + x = \bigcap_n O_n + x = \bigcap_n (O_n + x)$ is the intersection of open (hence measurable) sets, hence $A + x$ is also measurable. ☐

2.7 Countable Additivity, Continuity of Measure and Borel-Cantelli Lemma

This section is devoted to introducing important properties of outer measure. Some of them also hold for Lebesgue measure which will be soon redefined in another equivalent way. We will also prove that when $E \subseteq \mathbb{R}$ is measurable, then $m^*(E) = m(E)$ (where m^*, m are defined in the same way we did in Section 2.3), no matter bounded or not.

Theorem 2.7.1. *Outer measure is **countably additive** over measurable subsets, that is, if $\{E_k\}_{k=1}^{\infty}$ is a countable disjoint collection of measurable sets, then*

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

Proof. By subadditivity of outer measure we have clearly $m^*(\bigcup_{k=1}^{\infty} E_k) \leq \sum_{k=1}^{\infty} m^*(E_k)$. Moreover by Proposition 2.5.3 we deduce that for all $n \in \mathbb{N}$,

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

Definition 2.7.2. Let $\{E_k\}_{k=1}^{\infty}$ be a countable collection of subsets of $\mathbb{R}$.

(i) $\{E_k\}_{k=1}^{\infty}$ is **ascending** if $E_k \subseteq E_{k+1}$ for each k, and we define

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

(ii) $\{E_k\}_{k=1}^{\infty}$ is **descending** if $E_k \supseteq E_{k+1}$ for each k, and we define

$$\lim_{k\to\infty} E_k = \bigcap_{k=1}^{\infty} E_k.$$

Theorem 2.7.3 (Continuity of Measure). *Outer measure possesses the following properties:*

(i) If $\{A_k\}_{k=1}^{\infty}$ is an ascending collection of measurable sets, then

$$m^*\left(\bigcup_{k=1}^{\infty} A_k\right) := m^*\left(\lim_{k\to\infty} A_k\right) = \lim_{k\to\infty} m^*(A_k).$$

(ii) If $\{B_k\}_{k=1}^{\infty}$ is a descending collection of measurable sets and $m^*(B_N) < \infty$, for some $N \in \mathbb{N}$, then

$$m^*\left(\bigcap_{k=1}^{\infty} B_k\right) := m^*\left(\lim_{k\to\infty} B_k\right) = \lim_{k\to\infty} m^*(B_k).$$

Proof. (i) We have nothing to prove if one of $m^*(A_k) = \infty$. Let's assume $m^*(A_k) < \infty$ for all k and define $A_1' = A_1$, $A_k' = A_k \setminus A_{k-1}$ for $k \geq 2$, then $\bigcup_n A_n = \bigcup_n A_n'$ and by Theorem 2.7.1,

$$m^*\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_{n\to\infty} \left(\underbrace{m^*(A_1)}_{=m^*(A_1')} + \sum_{k=2}^{n} \underbrace{(m^*(A_k) - m^*(A_{k-1}))}_{=m^*(A_k')}\right) = \lim_{n\to\infty} m^*(A_n).$$

(ii) If there is an N such that $m(B_N) < \infty$, then we get an ascending collection of subsets $\{B_N \setminus B_k\}_{k>N}$, hence

$$m^*\left(\bigcup_{k=N+1}^{\infty} (B_N \setminus B_k)\right) = \lim_{k\to\infty} m^*(B_N \setminus B_k).$$

Since $\bigcup_{k=N+1}^{\infty}(B_N \setminus B_k) = B_N \setminus \bigcap_{k=N+1}^{\infty} B_k = B_N \setminus \bigcap_{k=1}^{\infty} B_k$, we are done by using the excision property of measurable sets. □

Proposition 2.7.4 (Outer Regularity of Lebesgue Measure). *If $E \subseteq \mathbb{R}$ is measurable, then $m^*(E) = m(E)$. Where m^* and m are both defined in Section 2.3.*

Proof. The only case that is unclear is when E is unbounded. For this case, let $E_n = E \cap [-n,n]$, then clearly E_n is bounded and measurable, $\bigcup_{n=1}^{\infty} E_n = E$ and $E_1 \subseteq E_2 \subseteq \cdots$. Hence by (i) of Theorem 2.7.3, one has

$$m^*(E) = m^*\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} m^*(E_n) = \lim_{n\to\infty} m(E_n) = m(E). \qquad \square$$

Remark. Lebesgue measure is also **inner regular**, i.e., $m_*(E) = m(E)$ for any measurable subset E of $\mathbb{R}$, which is left as exercise.

Definition 2.7.5. The **Lebesgue measure** is the set function m which is the restriction of outer measure m^* to the collection of measurable subsets of $\mathbb{R}$, i.e., $m = m^*|_{\mathcal{L}}$.

Remark.

- Since open set U is a disjoint union of open intervals, hence measurable, and thus by Proposition 2.3.4, $\lambda(U) = m^*(U) = m(U)$.
- For any closed L, L is measurable as its complement is open. Define $L_x = L \cap [-x,x]$ and let $U_x \supseteq L_x$ be open and have bounded measure, then $\lambda(L) := \lim_{x\to+\infty} \lambda(L_x) = \lim_{x\to+\infty} (\lambda(U_x) - \lambda(U_x \setminus L_x)) = \lim_{x\to+\infty} m(U_x \cap L_x) = \lim_{x\to+\infty} m(L_x) = m(L)$.

The above remark actually verifies that Lebesgue measure extends our definition of length.

Corollary 2.7.6. *First two theorems in this section are also true for Lebesgue measure m. That is, Lebesgue measure is **countably additive** and has the **continuity of measure** property.*

Proof. Since $m = m^*|_{\mathcal{L}}$, we replace m^* by m. □

Lemma 2.7.7 (Borel-Cantelli). *Let $\{E_k\}_{k=1}^{\infty}$ be a countable collection of measurable sets for which $\sum_{k=1}^{\infty} m(E_k) < \infty$. Then almost all $x \in \mathbb{R}$ belong to at most finitely many of the E_k's.*

Proof. First we observe that

$$A := \{x \in \mathbb{R} : x \text{ lies in infinitely many } E_k\text{'s}\} = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n.$$

We need to show $m(A) = 0$. Since $\sum_{k=1}^{\infty} m(E_k) < \infty$, by continuity of measure,

$$m(A) = m\left(\lim_{k\to\infty} \bigcup_{n=k}^{\infty} E_n\right) = \lim_{k\to\infty} m\left(\bigcup_{n=k}^{\infty} E_n\right) = 0. \qquad \square$$

2.8 Equivalence Relation

2.8.1 Brief Review of Equivalence Relation

Recall that an equivalence relation, $\sim$, on a set S is a binary relation that satisfies the following three conditions: **Reflexive:** For all $x \in S$, $x \sim x$. **Symmetric:** If $x \sim y$, then $y \sim x$. **Transitive:** If $x \sim y$, $y \sim z$, then $x \sim z$.

For $x \in S$, we introduce the equivalence class $[x] := \{s \in S : s \sim x\}$ and also $S/\sim$, the collection of such classes, i.e., $S/\sim = \{[s] : s \in S\}$, which is read as the quotient (set) of S by $\sim$. Recall that

$$a \sim b \iff [a] = [b] \quad \text{and} \quad a \not\sim b \iff [a] \neq [b] \iff [a] \cap [b] = \emptyset,$$

hence $\sim$ can be used to partition S, this is because $S = \bigcup_{s\in S}[s] = \bigsqcup_{\alpha\in A}[s_\alpha]$, where $s_\alpha \in S$ is called the **representative** of the class $[s_\alpha]$. Note that the feasibility of choosing those representatives follows from **Axiom of Choice**(2). The representative of a class may not be unique as we can choose (if exists) a $u_\alpha \in [s_\alpha] \setminus \{s_\alpha\}$ such that $u_\alpha \sim s_\alpha$ and thus $[u_\alpha] = [s_\alpha]$. Note that it is natural(3) to fix the representatives to avoid listing the same equivalence class.

For those who have had acquaintance with group theory the following example can be skipped, this is an example from number theory.

Example 2.8.1. Let $a, b \in \mathbb{Z}$, we can declare a relation $\sim$ on $\mathbb{Z}$ by

$$a \sim b \text{ if } a - b \in 2\mathbb{Z} := \{2n : n \in \mathbb{Z}\}.$$

For reflexivity, if $a \in \mathbb{Z}$, then $a - a = 0 \in 2\mathbb{Z}$.
For symmetry, if $a - b \in 2\mathbb{Z}$, then $b - a \in -2\mathbb{Z} = 2\mathbb{Z}$.
For transitivity, if $a - b \in 2\mathbb{Z}, b - c \in 2\mathbb{Z}$, then $a - c = (a - b) + (b - c) \in 2\mathbb{Z}$.
Let's compute $[a]$ when $a \in \mathbb{Z}$. By definition $[a] = \{n \in \mathbb{Z} : n \sim a\}$, so

$$[a] = \{n \in \mathbb{Z} : n - a = 2i, \exists i \in \mathbb{Z}\} = \bigcup_{i\in\mathbb{Z}} \{n \in \mathbb{Z} : n - a = 2i\} = \bigcup_{i\in\mathbb{Z}} \{a + 2i\} = a + 2\mathbb{Z}.$$

(2)Axiom of Choice is proved equivalent to Zorn's lemma whose application can be seen in chapter 0 of Kin Li's MATH371 notes (e.g. every vector space must have a basis).

(3)For example by fixing representatives one can show that any finite subgroups A, B of a group G must satisfy $|AB| = \frac{|A||B|}{|A\cap B|}$, where $AB := \{ab : a \in A, b \in B\}$.

Moreover by noting that $[a] = a + 2\mathbb{Z} = a - 2 + 2\mathbb{Z} = [a-2]$, it can be easily checked that
$$\mathbb{Z}/\!\sim\, = \{[n] : n \in \mathbb{Z}\} = \{[0],[1]\} = \big\{\{\text{even integers}\},\{\text{odd integers}\}\big\}. \qquad \square$$

2.8.2 Application of Equivalence Relation with a bit Group Theory

Given an equivalence relation $\sim$ on S, one can say that the collection in $S/\!\sim$ partitions S. One may also say that elements in an equivalence class $[s]$ are considered the same (or they are identified), which is a very common notion in mathematics (i.e., the concept of gluing or identifying things)! Let's elaborate this point with the help of group theory.

Definition 2.8.2. A **group** is a set G together with an associative operation $*$ such that

(i) If $a,b \in G$, then $a*b \in G$.

(ii) There is an element $e \in G$, called an **identity** of G, such that for all $g \in G$, $g*e = e*g = g$.

(iii) For each $g \in G$ there is an element $u \in G$ so that $g*u = u*g = e$.

Remark. It is a routine work to check the inverse of $g \in G$ is unique: Let $u,v \in G$ be the inverse of $g \in G$, then $u = u*e = u*(g*v) = (u*g)*v = e*v = v$. The inverse of g is denoted by g^{-1}.

Remark. When we say that G is a group, implicitly there is an operation between elements of G. It is customary to write ab instead of $a*b$, even though there may be two groups involved in a discussion. This convention is adopted as long as dropping "$*$" does not course any harm.

Definition 2.8.3. A set H is said to be a **subgroup** of a group G, denoted by $H \le G$, if it has the following properties:

(i) **Closure:** If $a,b \in H$, $ab \in H$.

(ii) **Identity:** $e \in H$.

(iii) **Inverses:** If $a \in H$, $a^{-1} \in H$.

Let G be a group and H its subgroup, we can check that the relation $\sim$ on G defined by $a \sim b$ if $b^{-1}a \in H$ (or equivalently, $a = bh$, for some $h \in H$) is an equivalence relation. The way we define $\sim$ is an analogue of Example 2.8.1. Now we see that
$$[a] = \{b \in G : b \sim a\} = \{b \in G : a^{-1}b \in H\} = \{b \in G : b \in aH\} = aH,$$
where aH is called a coset of G and $G/\!\sim\, = \{aH : a \in G\} =: G/H$ partitions G. By the notion of partition we easily check that $aH = bH$ iff $b^{-1}a \in H$, and in this case, we say that a and b are identified (in fact $[a] = [b]$, so a and b are "glued" together in the sense that they are squeezed into a set).

Since we can easily check an arbitrary intersection of groups is again a group, we can speak of finding the smallest subgroup H of G containing S by defining

$$H = \bigcap_{A \le G, A \supseteq S} A =: \langle S \rangle.$$

$\langle S \rangle$ is called a **subgroup of G generated by S**.

Having the preliminary knowledge, we try to find a subgroup H of a group G such that desired objects are "glued" in the sense that they lie in, or squeezed into, the same coset. For example, we just want to identify $a,b \in G$ and $c,d \in G$ respectively, so we hope after quotienting out G by H, $aH = bH, cH = dH$, which is the same as $b^{-1}a, d^{-1}c \in H$, so our group H must contain $b^{-1}a$ and $d^{-1}c$! Let $H = \langle b^{-1}a, d^{-1}c \rangle$, then of course $b^{-1}a, d^{-1}c \in H$, hence $aH = bH$ and $cH = dH$ in G/H, as desired. H is the **minimal** possible one to achieve this.

More generally, let $\mathcal{I}$ be an index set. For each $i \in \mathcal{I}$, we want to identify the elements in the subsets A_i and B_i respectively, then let H to be the group generated by their "difference"s

$$H = \langle b^{-1}a : a \in A_i, b \in B_i, i \in \mathcal{I} \rangle. \tag{2.8.4}$$

Of course we may take $H = G$, but the resulting quotient will be too trivial to be interested.

Example 2.8.5. It is obvious that $\mathbb{R}$ is a group under addition. Now we try to "identify" each $a \in [0,1]$ with $a \pm 1, a \pm 2, \ldots$, i.e., $a + k$, $k \in \mathbb{N}$. Then we need to quotient out $\mathbb{R}$ by a nice subgroup, what is that? By (2.8.4) to identify $A_a := \{a\}$ and $B_a := \{a + k : k \in \mathbb{Z}\}$, the subgroup needs to be generated by $-(a+k) + a = -k$, for all $k \in \mathbb{Z}$ and all $a \in [0,1]$. So we need

$$H = \langle -k : k \in \mathbb{Z}, a \in [0,1] \rangle = \langle \mathbb{Z} \rangle = \mathbb{Z},$$

hence $\mathbb{R}/\mathbb{Z}$ is the desired quotient. Geometrically, $\mathbb{R}/\mathbb{Z}$ is just the segment $[0,1)$ (a"="$a + k, k \in \mathbb{Z}$) with 0 and 1 identified, i.e., a circle depicted in Figure 2.1. By

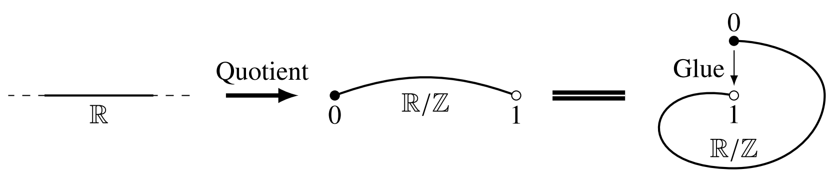


Figure 2.1: Quotient out $\mathbb{R}$ by $\mathbb{Z}$ to get S^1.

the same concept, it can be shown that $\mathbb{R}^2/\mathbb{Z}^2$ is a torus (by the way, $\mathbb{R}^n/\mathbb{Z}^n$ is called n-torus) as in Figure 2.2. $\square$

Definition 2.8.6. A group G is said to be **abelian** if for each $a,b \in G$, $ab = ba$.

Example 2.8.7. Let A,B be abelian groups (with $+$ denoting their binary operations) and let $f : A \to B$ be a set map. We want to quotient out B by a nice and smallest

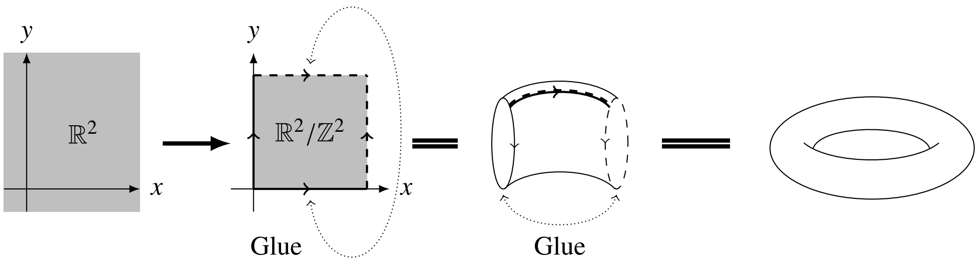


Figure 2.2: Quotient out $\mathbb{R}^2$ by $\mathbb{Z}^2$ to get a torus.

possible subgroup H such that the map

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[rr, bend right, "g := \pi \circ f"'] & B \arrow[r, "\pi"] & B/H \end{tikzcd}$$

preserves addition: $g(x+y) = g(x)+g(y)$. Here π is the canonical projection map. To do this, we want to identify $f(x+y)$ and $f(x)+f(y)$ for each $x,y \in A$, so we take

$$H = \langle f(x+y) - (f(x)+f(y)) : x,y \in A \rangle.$$

Then for any $x,y \in A$, $f(x+y)-(f(x)+f(y)) \in H$, so $f(x+y)+H = (f(x)+f(y))+H \iff g(x+y) = g(x)+g(y)$. $\square$

Example 2.8.8. Finally we use the smallest possible subgroup H of a group G to "abelianizes" G by doing quotient. Suppose G/H is "abelian", we hope that for each $a,b \in G$, $abH = baH$, that is, $a^{-1}b^{-1}ab \in H$, for all $a,b \in G$. This can be achieved if $H = \langle a^{-1}b^{-1}ab : a,b \in G \rangle$.

Surprisingly, H is normal in G (i.e., for each $g \in G$, $gHg^{-1} \subseteq H$), so the quotient G/H is again a group[4]. Since H is merely dependent on G, by defining $[a,b] = a^{-1}b^{-1}ab$, one uses the notation $[G,G]$ to denote such H (i.e., $[G,G] = \langle [a,b] : a,b \in G \rangle$) and call it a **commutator subgroup** of G. The abelianized group is usually denoted by $G/[G,G]$ (equipped with the quotient map $\pi : G \to G/[G,G]$) which is the following universal example: Given an abelian group A and a homomorphism[5] $\varphi : G \to A$, there is a unique homomorphism $\tilde{\varphi} : G/[G,G] \to A$ such that $\varphi = \tilde{\varphi} \circ \pi$. That said, the following diagram commutes.

$$\begin{tikzcd} G \arrow[r, "\pi"] \arrow[dr, "\varphi"'] & G/[G,G] \arrow[d, dotted, "\tilde{\varphi}"] \\ & A \end{tikzcd}$$

[4] We don't go into detail, the resulting group is called a quotient group which is usually mentioned in abstract algebra text.

[5] i.e., a map ϕ between groups such that $\phi(ab) = \phi(a)\phi(b)$.

We also say that φ factors through $G/[G,G]$. □

2.9 Nonmeasurable Sets

Theorem 2.9.1. *Any subset E of $\mathbb{R}$ with positive outer measure contains a subset that fails to be measurable.*

Proof. By Problem 2.2 we can assume E is bounded, otherwise consider its bounded subset with positive outer measure. Now we define $x \sim y$ on E if $x - y \in \mathbb{Q}$, it is easy to check $\sim$ is an equivalence relation. We partition E by $E/{\sim} = \{[c] : c \in C_E\}$, where $E = \bigsqcup_{c \in C_E} [c]$ and $C_E \subseteq E$, here C_E is also bounded.

For the sake of contradiction let's assume C_E is measurable. We first prove that $m(C_E) = 0$. Let $Q = \mathbb{Q} \cap [0,1]$, then $\bigcup_{q \in Q}(q + C_E)$ is bounded. Moreover, the collection $\{q + C_E\}_{q \in Q}$ is disjoint because if there are $q_1, q_2 \in Q$ such that $(q_1 + C_E) \cap (q_2 + C_E) \neq \emptyset$, then there are $c_1, c_2 \in C_E$ such that $q_1 + c_1 = q_2 + c_2 \implies c_1 - c_2 \in \mathbb{Q} \implies [c_1] = [c_2] \implies c_1 = c_2$, and thus $q_1 = q_2$. Now by countable additivity of Lebesgue measure (recall Example 2.6.2),

$$m\left(\bigsqcup_{q \in Q}(q + C_E)\right) = \sum_{q \in Q} m(q + C_E) < \infty,$$

however, $m(q + C_E) = m(C_E)$ (by Proposition 2.3.5 and Proposition 2.7.4), forcing $m(C_E) = 0$.

Since $E = \bigsqcup_{c \in C_E} [c]$, if $x \in E$, there is a $c \in C_E$ such that $x \in [c]$, this implies there is $q \in \mathbb{Q}$ such that $x = q + c \in q + C_E$, meaning the inclusion

$$E \subseteq \bigcup_{q \in \mathbb{Q}} (q + C_E),$$

it follows from subadditivity and translation invariance property of outer measure that

$$m^*(E) \le \sum_{q \in \mathbb{Q}} m^*(q + C_E) = \sum_{q \in \mathbb{Q}} m^*(C_E) = 0,$$

a contradiction □

2.10 Further Topic

This section is devoted to the preparation of general measure space. Acquaintence with more kinds of σ-algebra arguably helps create counter examples to show theory on Lebesgue measure can fail in general measure.

The Cantor-Lebesgue function constructed in this section not only shows us Borel σ-algebra is properly contained in the σ-algebra of Lebesgue measurable sets, but also provides us a concrete example that the composition of measurable functions (will be defined in next chapter) needs not be measurable.

Finally we end this chapter by providing two propositions which allows us to understand the structure of measurable sets geometrically.

2.10.1 Borel sigma-algebra

Definition 2.10.1. Let X be a set. A $\boldsymbol{\sigma}$**-algebra** on X is a collection $\mathcal{S} \subseteq 2^X$ satisfying the following properties:

(i) $X \in \mathcal{S}$.

(ii) If $E \in \mathcal{S}$, then $X \setminus E \in \mathcal{S}$.

(iii) If $E_1, E_2, \cdots \in \mathcal{S}$, then $\bigcup_{i=1}^{\infty} E_i \in \mathcal{S}$.

Remark. (ii) and (iii) of Definition 2.10.1 imply σ-algebra is closed under countable intersection.

Up til now the only σ-algebra we have discussed is the collection of Lebesgue measurable subsets of $\mathbb{R}$, $\mathcal{L}$, in Section 2.5. What is any other else on $\mathbb{R}$?

Example 2.10.2 (A list of some σ-algebras on $\mathbb{R}$, S).

(i) $S = \{\emptyset, \mathbb{R}\}$, called **trivial $\boldsymbol{\sigma}$-algebra**.

(ii) $S = 2^{\mathbb{R}}$, the collection of all subsets of $\mathbb{R}$.

(iii) If $E \subseteq \mathbb{R}$ is a proper subset, $S = \{\emptyset, E, \mathbb{R} \setminus E, \mathbb{R}\}$.

(iv) Let $X \subseteq \mathbb{R}$ be uncountable, $S = \{A \subseteq X : A \text{ countable or } X \setminus A \text{ countale}\}$. $\square$

In particular, the σ-algebra generated by the collection of open intervals of $\mathbb{R}$ in the way of Definition 2.10.1 is called the **Borel σ-algebra**, denoted by $\mathcal{B}$, which is the smallest σ-algebra that contains all the open intervals (we call talk about the smallest one, thanks to Problem 2.17). We call any element of $\mathcal{B}$ a Borel measurable set, or simply a Borel set.

Different people may use topologically different "generator"s, the following is for reference and its proof is tedious and thus omitted.

Proposition 2.10.3. *The Borel σ-algebra of subsets of $\mathbb{R}$, $\mathcal{B}$ may be decribed as the σ-algebra generated by these families of subets of $\mathbb{R}$:*

(i) Open intervals.

(ii) Open sets.

(iii) Closed intervals.

(iv) Closed sets.

(v) Compact sets.

(vi) Left open, right closed intervals.

(vii) Left closed, right open intervals.

(viii) All intervals.

Clearly $\mathcal{B} \subseteq \mathcal{L}$ because $\mathcal{L}$ contains all open intervals and is itself a σ-algebra. However, is $\mathcal{B}$ a proper subset of $\mathcal{L}$? The answer is positive, we will prove the existence of nonBorel measurable set after the construction of Cantor-Lebesgue function in next subsection.

2.10.2 Cantor-Lebesgue Function

In this subsection we will recall what is Cantor set, our construction of Cantor-Lebesgue function will be first defined on the complement of Cantor set and will then be extended to all of $[0,1]$.

To get the structure of Cantor set straight, it is helpful to get the feeling form its complement first. Let C_n be the n^{th} stage of the construction of Cantor set, they are loosely depicted in Figure 2.3.

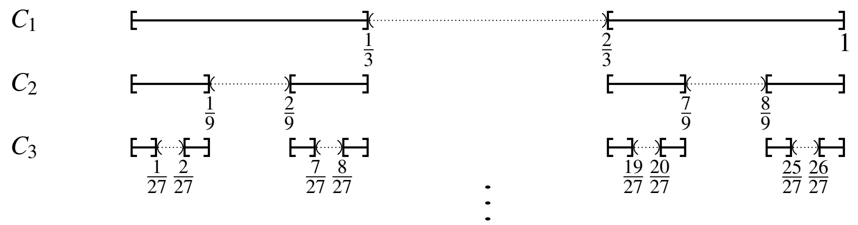


Figure 2.3: Cantor set.

In each step we divide each existing closed interval into three pieces evenly and remove the middle one, with end points left there. Now the collection $\{C_n\}$ is descending, we define **Cantor set** to be the limit of this collection. That is,

$$C := \lim_{n\to\infty} C_n := \bigcap_{n=1}^{\infty} C_n.$$

Observe that $m(C_n) = 2^n \times \frac{1}{3^n}$, and then by continuity of measure,

$$m(C) = m\left(\lim_{n\to\infty} C_n\right) = \lim_{n\to\infty} m(C_n) = \lim_{n\to\infty}\left(\frac{2}{3}\right)^n = 0.$$

Moreover, C is clearly closed and as C has one-one correspondence with $\{0,1\}\times\{0,1\}\times\cdots$, it is uncountable. That said, a set of measure zero is not necessarily countable.

Let O_n be the open set that is removed in the first n stages of the construction of Cantor set. That is,

$$O_n = [0,1]\setminus C_n.$$

We define $O = \bigcup_{n=1}^{\infty} O_n$, clearly O is open in $\mathbb{R}$, dense in $[0,1]$ with $m(O) = 1$. We see O_k contains $2^k - 1$ disjoint open intervals. For detailed discussion we let $O_k = \bigsqcup_{i=1}^{2^k-1} I_i^k$, where I_i^k denotes the i^{th} open interval of O_k counted from the left.

Now we are ready to construct our **Cantor-Lebesgue function** $\varphi : [0,1] \to \mathbb{R}$. For each $k \in \mathbb{N}$, we define

$$\varphi|_{O_k} = \sum_{i=1}^{2^k-1} \frac{i}{2^k}\chi_{I_i^k},$$

where for $A \subseteq \mathbb{R}$, $\chi_A(x)$ is 1 when $x \in A$ and 0 when $x \notin A$. $\varphi|_{O_k}$ is constant on each I_j^k and takes the values $\frac{1}{2^k}, \frac{2}{2^k}, \ldots, \frac{2^k-1}{2^k}$ increasingly. For example,

$$\begin{aligned}
\varphi|_{O_1} &= \frac{1}{2}\chi_{I_1^1} \\
\varphi|_{O_2} &= \frac{1}{4}\chi_{I_1^2} + \frac{1}{2}\chi_{I_1^1} + \frac{3}{4}\chi_{I_3^2}
\end{aligned}$$

$$\varphi|_{O_3} = \frac{1}{8}\chi_{I_1^3} + \frac{1}{4}\chi_{I_1^2} + \frac{3}{8}\chi_{I_3^3} + \frac{1}{2}\chi_{I_1^1} + \frac{5}{8}\chi_{I_5^3} + \frac{3}{4}\chi_{I_3^2} + \frac{7}{8}\chi_{I_7^3}.$$

It can be seen that $\varphi|_{O_{k+1}}$ extends $\varphi|_{O_k}$, $k \geq 1$.

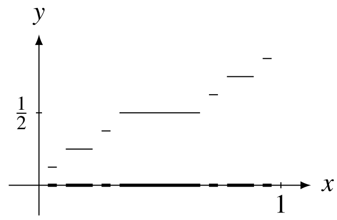


Figure 2.4: Construction in the third stage: $\varphi|_{O_3}$.

We have thereby defined φ on all of O. Now define $\varphi(0) = 0$ and $\varphi(x) = \sup f(O \cap [0,x))$ for nonzero $x \in [0,1] \setminus O = C$ (it is automatic that $\varphi(1) = 1$).

Proposition 2.10.4. *Cantor-Lebesgue function φ has the following properties:*

(i) Increasing on $[0,1]$; *(ii) Continuous on* $[0,1]$.

Proof. (i) follows from the observation that $\varphi(x) = \sup \varphi(O \cap [0,x))$ for all $x \in O$.

(ii) φ is clearly continuous on O. Let $x \in [0.1] \setminus O$, as the only discontinuity of an increasing function is a "jump", hence it is enough to show φ can't have a jump at x. Let's first assume $x \neq 0,1$. Since O is dense in $[0,1]$, there must be $a,b \in O$ such that $a < x < b$. But then there is an $N \in \mathbb{N}$ so that $a,b \in O_N$. By the ascending property of $\{O_n\}$, for each $n \geq N$ there must be an index $k_n \in \mathbb{N}$ such that x lies between $I_{k_n}^n$ and $I_{k_n+1}^n$. Choose $a_n = \sup I_{k_n}^n$ and $b_n = \inf I_{k_n+1}^n$, we see that

$$a_n \leq x \leq b_n, \quad b_n - a_n = \frac{1}{3^n} \quad \text{and} \quad \varphi(b_n) - \varphi(a_n) = \frac{1}{2^n}.$$

The process can always continue whenever $n \geq N$, thus x cannot be a jump discontinuity. The continuity at 0 and 1 can be similarly proved. □

Remark. φ maps $[0,1]$ onto $[0,1]$ by intermediate value theorem.

Proposition 2.10.5. *Construct the strictly increasing continuous function ψ on* $[0,1]$ *as follows:*

$$\psi(x) = x + \varphi(x).$$

Where φ is the Cantor-Lebesgue function, then ψ has the following properties:

(i) ψ maps the Cantor set onto a set of positive measure.

(ii) ψ maps a subset of Cantor set onto a nonmeasurable set.

Proof. (i) Observe that φ is constant on each I_j^k and thus ψ takes I_j^k onto a segment with the same length, so $m(\psi(O_k)) = m(O_k) \implies m(\psi(O)) = m(O) = 1$. On the other hand, since ψ maps $[0,1]$ onto $[0,2]$,

$$2 = m([0,2]) = m(\psi(O) \sqcup \psi(C)) = 1 + m(\psi(C)) \implies m(\psi(C)) = 1 > 0.$$

(ii) By Theorem 2.9.1, that $\psi(C)$ has positive measure implies there is a nonmeasurable $T \subseteq \psi(C)$, hence $N := \psi^{-1}(T) \subseteq C$ is mapped onto $\psi(N) = T$. □

Remark. In contrast to Problem 2.7 a continuous function does not necessarily take a set of measure zero to measure zero.

Proposition 2.10.6. *If f is a continuous strictly increasing function of $\mathbb{R}$ onto $\mathbb{R}$, then f maps Borel set to Borel set.*

Proof. It is clear that f takes compact interval to compact interval, thus it suffices to show that

$$\mathcal{S} := \{A \subseteq \mathbb{R} : f(A) \text{ is Borel}\}$$

is a σ-algebra on $\mathbb{R}$ since it already contains all the "generator"s. The detail is left as exercise. □

Theorem 2.10.7. *There is a subset of Cantor set that is Lebesgue measurable but not Borel.*

Proof. Extend ψ in Proposition 2.10.5 to Ψ on $\mathbb{R}$ increasingly with constant positive slope outside $[0,1]$, then Ψ becomes a continuous strictly increasing function from $\mathbb{R}$ onto $\mathbb{R}$. The Lebesgue measurable N in part (ii) of the proof of Proposition 2.10.5 can't be Borel, otherwise the nonmeasurable $\Psi(N) = \psi(N)$ must also be Borel by Proposition 2.10.6, a contradiction. □

2.10.3 Geometric Structure of Measurable Sets

There are still two facts that are not hard to understand and worth seeing once. They provide us with a geometric view how a measurable set of positive measure looks.

Lemma 2.10.8. *Let $E \subseteq \mathbb{R}$ be measurable with $m(E) > 0$, then for any $\lambda \in (0,1)$ there is an open interval I such that*

$$\lambda m(I) < m(I \cap E).$$

That said, measurable sets with positive measure can always be "squeezed" into an **bounded**(6) open interval and "λ" acts as a squeezing factor.

Proof. Let's assume $m(E) < \infty$, otherwise consider its bounded subset. Let $\lambda \in (0,1)$ be given, then for each $\epsilon > 0$ one can find an open $U_\epsilon \supseteq E$ such that $m(U_\epsilon) < m(E) + \epsilon$. Let $U_\epsilon = \bigsqcup_i I_i^\epsilon$, we claim that one of I_i^ϵ's satisfies our desired inequality. Suppose not, then for all i,

$$\lambda m(I_i^\epsilon) \geq m(I_i^\epsilon \cap E) \implies \lambda \sum_i m(I_i^\epsilon) \geq \sum_i m(I_i^\epsilon \cap E) = m(U_\epsilon \cap E) = m(E),$$

and hence

$$\lambda(m(E) + \epsilon) > m(E).$$

Since $\lambda \in (0,1)$, it is straightforward to see when ϵ is too small, we get a contradiction that $m(E) > m(E)$. This contradiction arises whenever $\epsilon < m(E)(\frac{1}{\lambda} - 1)$, so by taking $\epsilon = 0.9999 \cdot m(E)(\frac{1}{\lambda} - 1)$ at the beginning, we are done. □

(6)This is implicit in the inequality of the lemma as the case $m(I) = \infty$ does not make sense.

Theorem 2.10.9 (Steinhaus). *Let $E \subseteq \mathbb{R}$ be measurable with $m(E) > 0$, let*

$$E - E = \{x - y : x, y \in E\},$$

then there is a $\delta > 0$ such that $E - E \supseteq B(0,\delta)$.

In other words, 0 must be an interior point of $E - E$ if E has positive measure.

Proof. Let $\lambda \in (\frac{3}{4}, 1)$ be given[7], then by the lemma above one can find a "container" (an open interval) I such that E is partially squeezed into I satisfying

$$m(E \cap I) > \lambda m(I).$$

Let's notice that when $x_0 \in B(0,\delta)$, $x_0 \in E - E \iff (E + x_0) \cap E \neq \emptyset$. On account of the last inequality, it is more preferable to consider the squeezed one (as we know more about it). In other words, our goal is to choose δ small such that $(E \cap I + x_0) \cap (E \cap I) \neq \emptyset$.

As $I = B(a,r)$, let's construct $J = B(0, \frac{r}{2})$. Let $x_0 \in J$, then $m\big((I + x_0) \cap I\big) > \frac{1}{2} m(I)$, it follows that

$$m\big((I + x_0) \cup I\big) = 2m(I) - m\big((I + x_0) \cap I\big) < \tfrac{3}{2} m(I).$$

We claim that $(E \cap I + x_0) \cap (E \cap I) \neq \emptyset$, suppose not,

$$\begin{aligned}
\tfrac{3}{2} m(I) &> m\big((I + x_0) \cup I\big) \\
&\geq m\big((E \cap I + x_0) \cup (E \cap I)\big) \\
&= m(E \cap I + x_0) + m(E \cap I) \\
&> 2 \times \lambda m(I) \\
&> 2 \times \tfrac{3}{4} m(I) \\
&= \tfrac{3}{2} m(I),
\end{aligned}$$

a contradiction. □

2.11 Exercises and Problems

Exercises

2.1. Suppose A and B differ by a set of measure zero. In other words,

$$m^*\big((A \setminus B) \cup (B \setminus A)\big) = 0.$$

Prove that A is measurable if and only if B is measurable. Moreover, we have $m(A) = m(B)$ in case both are measurable.

2.2. Show that if a set E has positive outer measure, then there is a bounded subset of E that has positive outer measure.

2.3. In the text there are two (why?)'s, one is in the proof of Proposition 2.3.2 and one is in the proof of Lemma 2.4.3, they are as shown below, explain.

Let $\{K_i\}_{i=1}^n$ be a finite collection of compact subsets of $\mathbb{R}$, show that:

[7]We will see why we take it that way in the -4 line of the proof.

(a) $\lambda(K_1) \le \lambda(K_1 \cup K_2)$.

(b) If $K_i \cap K_j = \emptyset$ for $i \neq j$, $\sum_{i=1}^n \lambda(K_i) = \lambda(\bigcup_{i=1}^n K_i)$.

2.4. For a collection of sets $\{E_i\}$ we have shown that $m^*(\bigcup E_i) \le \sum m^*(E_i)$ in Proposition 2.3.2. Show that if a collection of sets $\{A_i\}$ is disjoint, then

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

2.5. Prove the following:

(a) Show that if E,F are measurable, then $m(E \cup F) + m(E \cap F) = m(E) + m(F)$.

(b) Consider $E \subseteq \mathbb{R}$. Show that there is a G_δ set $G \supseteq E$ such that $m(G) = m^*(E)$.

(c) Let $m^*(A \cup B) < \infty$, show that if $m^*(A \cup B) = m^*(A) + m^*(B)$, then $A \cap B$ is measurable.

Problems

2.6. Let $f : X \to \mathbb{R}$ be L-Lipschitz, i.e., there is a constant L such that for all $x,y \in X$, $|f(x) - f(y)| \le L|x-y|$. Show that the function

$$F(x) := \inf_{a \in X} \big(f(a) + L|x-a|\big)$$

is Lipschitz on $\mathbb{R}$ and extends f.

2.7. Let $f : X \to \mathbb{R}$ be L-Lipschitz, where $X \subseteq \mathbb{R}$. That is, there is a constant L such that $|f(x) - f(y)| \le L|x-y|$ for any $x,y \in X$. Show that for any $A \subseteq X$, one has

$$m^*\big(f(A)\big) \le Lm^*(A).$$

Then prove that a Lipschitz function takes bounded measurable sets to bounded measurable sets. Prove also that it takes any measurable set to measurable set.
[**Hint:** For measurability you may use Theorem 2.6.1.]

2.8. Prove that for any measurable $A \subseteq \mathbb{R}$, one has $m(A) = m_*(A)$.

2.9. Let $E \subseteq \mathbb{R}$. If for each $x \in E$, there is an open interval $(x - \delta_x, x + \delta_x)$ such that

$$m^*\big(E \cap (x - \delta_x, x + \delta_x)\big) = 0,$$

prove that $m^*(E) = 0$.

2.10. Let E have finite outer measure. Show that E is measurable if and only if for each open and bounded interval (a,b),

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

2.11. Suppose f and g are continuous functions on $[a,b]$. Show that if $f = g$ a.e. on $[a,b]$, then, in fact, $f = g$ on $[a,b]$. Is a similar assertion true if $[a,b]$ is replaced by a general measurable set E?

2.12. (Dini's theorem) Let $\{f_n\}$ be an increasing sequence of continuous functions on $[a,b]$ which converges pointwise on $[a,b]$ to the continuous function f on $[a,b]$. Show that the convergence is uniform on $[a,b]$.
[**Hint:** For $\epsilon > 0$ and for each natural number n, show that $\{E_n\}$ defined by $E_n = \{x \in [a,b] : f(x) - f_n(x) < \epsilon\}$ is an open cover of $[a,b]$.]

2.13. (MATH301 1998 Final) Write the rational numbers $\mathbb{Q} = \{q_1, q_2, q_3, \dots\}$. Define

$$G = \bigcup_{n=1}^{\infty} \left(q_n - \frac{1}{n^2}, q_n + \frac{1}{n^2} \right).$$

(a) Show that G is measurable and $m(G) < \infty$.

(b) Show that if F is a closed set and $\mathbb{Q} \subseteq F$, then $F = \mathbb{R}$.

(c) For every closed set F, show that $m(G \setminus F) > 0$ or $m(F \setminus G) > 0$.

2.14. Show that $E \subseteq \mathbb{R}$ has measure zero if and only if there is a countable collection of open intervals $\{I_k\}_{k=1}^{\infty}$ for which each point in E belongs to infinitely many of the I_k's and $\sum_{k=1}^{\infty} \lambda(I_k) < \infty$.

2.15. (Riesz-Nagy) Let E be a set of measure zero contained in the open interval (a,b). According to the Problem 2.14, there is a countable collection of open intervals contained in (a,b), $\{(c_k,d_k)\}_{k=1}^{\infty}$, for which each point in E belongs to infinitely many intervals in the collection and $\sum_{k=1}^{\infty}(d_k - c_k) < \infty$. Define

$$f(x) = \sum_{k=1}^{\infty} \lambda\big((c_k,d_k) \cap (-\infty,x)\big)$$

for all $x \in (a,b)$. Show that f is increasing and fails to be differentiable at each point in E.

2.16. Let E be a measurable subset of $\mathbb{R}$, $m(E) < \infty$ and $\{f_n\}$ be a sequence of measurable functions on E. Let $\{\alpha_n\}$ be a sequence of positive numbers such that $\sum_{n=1}^{\infty} m\{x \in E : |f_n(x)| > \alpha_n\} < \infty$. Prove that

$$-1 \leq \varliminf_{n\to\infty} \frac{f_n(x)}{\alpha_n} \leq \varlimsup_{n\to\infty} \frac{f_n(x)}{\alpha_n} \leq 1$$

for almost all $x \in E$.

2.17. Prove that the intersection of σ-algebras is a σ-algebra. In particular, for any collection of subsets, we may talk about the smallest σ-algebra that contains the collection.

2.18. Show that the 4$^{\text{th}}$ one in Example 2.10.2 is a σ-algebra. Also complete the proof of Proposition 2.10.6, that is, check that $\mathcal{S}$ is a σ-algebra.

2.19. Is there a measurable set E such that for any $(a,b) \subseteq [0,1]$, $m(E \cap (a,b)) = \dfrac{b-a}{2}$?

2.20. (MATH301 2003 Final) Let E be a bounded measurable set in $\mathbb{R}$ such that $m(E \cap I) \leq \frac{1}{2} m(I)$ for every interval I. Prove that $m(E) = 0$.

Chapter 3

Lebesgue Measurable Functions

This chapter is devoted to the study of measurable functions which lays down the foundation of Lebesgue integration. For example, the natural objects like continuous functions, monotone functions and step functions are all "measurable" (to be defined). We will also establish results concerning the approximation of measurable functions by simple functions and continuous functions.

3.1 Sums, Products and Compositions

To avoid notations being cumbersome, we will write $f^{-1}\langle a,b\rangle$ instead of $f^{-1}\big(\langle a,b\rangle\big)$, $f^{-1}(c)$ instead of $f^{-1}\big(\{c\}\big)$ and $m\{x:P\}$ instead of $m\big(\{x:P\}\big)$.

Proposition 3.1.1. *Let the function f have a measurable domain E. Then the following statements are equivalent.*

(i) For each $c\in\mathbb{R}$, $\{x\in E: f(x)>c\}$ is measurable.

(ii) For each $c\in\mathbb{R}$, $\{x\in E: f(x)\geq c\}$ is measurable.

(iii) For each $c\in\mathbb{R}$, $\{x\in E: f(x)<c\}$ is measurable.

(iv) For each $c\in\mathbb{R}$, $\{x\in E: f(x)\leq c\}$ is measurable.

Proof. $f(x)\geq c \iff f(x)>c-\frac{1}{n},\forall n\in\mathbb{N}$ and $f(x)>c \iff f(x)\geq c+\frac{1}{n},\exists n\in\mathbb{N}$, we see that

$$\{x\in E: f(x)\geq c\}=\bigcap_{n=1}^{\infty}\left\{x\in E: f(x)>c-\frac{1}{n}\right\},$$

$$\{x\in E: f(x)>c\}=\bigcup_{n=1}^{\infty}\left\{x\in E: f(x)\geq c+\frac{1}{n}\right\},$$

hence (i) $\Leftrightarrow$ (ii), the proof that (iii) $\Leftrightarrow$ (iv) is essentially the same. That (ii) $\Leftrightarrow$ (iii) is obvious. $\square$

Corollary 3.1.2. *With the same hypothesis in Proposition 3.1.1 and if one of the 4 statements holds, then for each extended real number* c*, the set* $\{x \in E : f(x) = c\} = f^{-1}(c)$ *is measurable.*

Proof. When $|c| < \infty$, $\{c\} = \bigcap_{n=1}^{\infty}(c - \frac{1}{n}, c + \frac{1}{n})$, it follows easily from the preceding proposition that

$$f^{-1}(c) = \begin{cases} \displaystyle\bigcap_{n=1}^{\infty} f^{-1}\left(c - \frac{1}{n}, c + \frac{1}{n}\right), & \text{if } |c| < \infty, \\ \displaystyle\bigcap_{n=1}^{\infty} \{x \in E : f(x) > n\}, & \text{if } c = +\infty, \\ \displaystyle\bigcap_{n=1}^{\infty} \{x \in E : f(x) < -n\}, & \text{if } c = -\infty \end{cases}$$

is measurable. □

Usually we would like to partition the range of a “nice” function f into (small) intervals whose pre-image is expected be measurable such that we can approximate f by some “simple” functions for which we can define an integral analogous to Riemann integral.

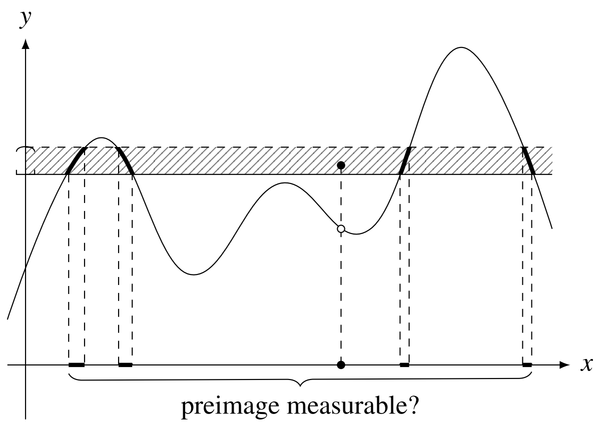


Figure 3.1: Preimage of measurable function.

Proposition 3.1.1 tells us one of the conditions suffices to show f is “nice” because, for example, $f^{-1}[a, b) = f^{-1}[a, +\infty) \cap f^{-1}(-\infty, b)$. We call those “nice” functions *measurable*. We retain the adjective *simple* to describe such “simple” functions in Definition 3.3.1.

More often instead of real-valued function we are interested in extended real-valued function. That is, a function that not only takes the value in $\mathbb{R}$, but also $-\infty$ or $+\infty$ with the arithmetic defined as follows:

$$\begin{aligned} a + \infty = +\infty + a &= +\infty, & a &\neq -\infty \\ a - \infty = -\infty + a &= -\infty, & a &\neq +\infty \\ a \cdot (\pm\infty) = \pm\infty \cdot a &= \pm\infty, & a &\in (0, +\infty] \end{aligned}$$

$$
\begin{aligned}
a \cdot (\pm\infty) = \pm\infty \cdot a = \mp\infty, \qquad & a \in [-\infty, 0) \\
\frac{a}{\pm\infty} = 0, \qquad & a \in \mathbb{R} \\
\frac{\pm\infty}{a} = \pm\infty, \qquad & a \in (0, +\infty) \\
\frac{\pm\infty}{a} = \mp\infty, \qquad & a \in (-\infty, 0)
\end{aligned}
$$

We also adopt the convention that $0 \cdot (\pm\infty) = 0$. We can denote the extended real line by $[-\infty, \infty]$, $\hat{\mathbb{R}}$ or $\mathbb{R}^*$, some may also denote it by $\overline{\mathbb{R}}$ but it repeats the already defined set operation $\overline{\cdot}$ (taking closure) for which $\overline{\mathbb{R}} = \mathbb{R}$.

Handling the "number" ∞ provides us a higher generality, this is convenience since, for example, no matter how bad the function $\sum_{n=1}^\infty f_n(x)$ behaves on E, as long as $m(E) = 0$, it is still manageable to ignore E in application.

Definition 3.1.3. A function f defined on E is said to be **Lebesgue measurable**, or simply **measurable**, provided it is extended real-valued, its domain E is measurable and it satisfies one of the four statements of Proposition 3.1.1.

Remark. The domain of a measurable function is tacitly assumed measurable. To construct a nonmeasurable function we usually define it on a measurable domain and argue one of the conditions in Proposition 3.1.1 cannot hold.

Proposition 3.1.4. *Let the function f be defined on a measurable set E. Then f is measurable if and only if $f^{-1}(O)$ is measurable for each open set O.*

Proof. Assume f is measurable. Let O be open, $O = \bigsqcup (a_i, b_i)$, where $a_i, b_i \in [-\infty, \infty]$, then $f^{-1}(O) = \bigsqcup f^{-1}(a_i, b_i)$ is measurable since $(a_i, b_i) = (a_i, \infty) \cap (-\infty, b_i)$.

Conversely, for any $a \in \mathbb{R}$, as (a, ∞) is open, $f^{-1}(a, \infty)$ is measurable. Then since $f^{-1}(+\infty) = \bigcap_{n=1}^\infty f^{-1}(n, \infty)$, we conclude for each $c \in \mathbb{R}$, $\{x \in E : f(x) > c\}$ is measurable. □

Proposition 3.1.5. *A real-valued function that is continuous on its measurable domain E is measurable.*

Proof. Since f is continuous for each open O, there is an open $U \subseteq \mathbb{R}$ such that $f^{-1}(O) = U \cap E$, an intersection of two measurable sets. So from Proposition 3.1.4 f is measurable. □

Proposition 3.1.6. *A monotone function that is defined on an interval is measurable.*

Proof. We leave it as an exercise. □

In measure theory sets of measure zero are considered to be "negligible" sets in many sense. This point will be made clear in the study of measurable function and integration. We are interested in whether a function possesses certain properties *except a set of measure zero*. This is closely related to the following concept:

Definition 3.1.7. Let $E \subseteq \mathbb{R}$ be measurable and let $P(x)$ be a property related to points $x \in \mathbb{R}$. If $m\{x \in E : P(x) \text{ does not hold}\} = 0$, we say that $P(x)$ holds **almost everywhere (abbr. a.e.) on E**, or that $P(x)$ holds for **almost every (abbr. a.e.) $x \in E$**.

Remark. If E has finite measure, then

$$P(x) \text{ holds a.e. on } E \iff m\{x \in E : P(x) \text{ holds}\} = m(E),$$

this is because $m(E \setminus \{x \in E : P(x) \text{ holds}\}) = m(E) - m\{x \in E : P(x) \text{ holds}\}$.

Proposition 3.1.8. *Let f be an extended real-valued function on E.*

(i) If f is measurable on E and $f = g$ a.e. on E, then g is measurable on E.

(ii) For a measurable subset D of E, f is measurable on E if and only if the restrictions of f to D and $E \setminus D$ are measurable.

Proof. (i) Let $E_0 \subseteq E$ be such that $f = g$ on $E \setminus E_0$ and $m(E_0) = 0$. Let $c \in \mathbb{R}$, then

$$\{x \in E : g(x) > c\} = \{x \in E \setminus E_0 : g(x) > c\} \sqcup \{x \in E_0 : g(x) > c\}. \tag{3.1.9}$$

The measuability follows by seeing $\{x \in E \setminus E_0 : g(x) > c\} = (E \setminus E_0) \cap \{x \in E : f(x) > c\}$ and $m\{x \in E_0 : g(x) > c\} = 0$.

(ii) Just observe that

$$\{x \in E : f(x) > c\} = \{x \in D : f(x) > c\} \sqcup \{x \in E \setminus D : f(x) > c\}. \qquad \square$$

For two measurable functions on a common domain E, the sum $f + g$ may not be properly defined when one takes the value $+\infty$ while another one takes $-\infty$. But if they are finite a.e. on E, then there is an $E_0 \subseteq E$ such that f, g are finite on $E \setminus E_0$ but $m(E_0) = 0$.

Theorem 3.1.10. *Let f and g be measurable functions on E that are finite a.e. on E.*

(i) For any $\alpha, \beta \in \mathbb{R}$, $\alpha f + \beta g$ is measurable on E.

(ii) $f \cdot g$ is measurable on E.

Proof. Let $E_0 \subseteq E$ be such that $m(E_0) = 0$ and f and g are finite on $E \setminus E_0$.

(i) If $\alpha = 0$, then αf is clearly measurable. If $\alpha \neq 0$, then $\{x \in E \setminus E_0 : \alpha f(x) > c\} = \{x \in E \setminus E_0 : f(x) \lesseqgtr c/\alpha\}$ is measurable by (ii) of Proposition 3.1.8, i.e., αf is measurable. It suffices to consider the case $\alpha = \beta = 1$.

Now for each $x \in E$, $f(x) + g(x) > c \iff f(x) > c - g(x) \iff f(x) > r > c - g(x)$ for some $r \in \mathbb{Q}$, hence

$$\begin{aligned}
&\{x \in E \setminus E_0 : f(x) + g(x) > c\} \\
&= \{x \in E \setminus E_0 : f(x) > r > c - g(x), \exists r \in \mathbb{Q}\} \\
&= \bigcup_{r \in \mathbb{Q}} [\{x \in E \setminus E_0 : f(x) > r\} \cap \{x \in E \setminus E_0 : r > c - g(x)\}],
\end{aligned}$$

therefore the measurability of $f + g$ is clear.

(ii) The identity $f \cdot g = \frac{1}{2}\big((f+g)^2 - f^2 - g^2\big)$ tells us it suffices to study the measurability of f^2. It just takes some time to fill up the detail. $\square$

Proposition 3.1.11. *Let g be a continuous real-valued function defined on all of $\mathbb{R}$ and f a measurable real-valued function defined on E. Then the composition $g \circ f$ is measurable on E.*

Proof. We use Proposition 3.1.4. Let O be open, then $(g\circ f)^{-1}(O) = f^{-1}(g^{-1}(O))$, but $g^{-1}(O)$ is open in $\mathbb{R}$, $f^{-1}(g^{-1}(O))$ is then measurable by measurability of f. □

Remark. The above proposition can be false if **(i)** g is merely measurable; **(ii)** g is measurable and f is continuous. See Problem 3.5.

Proposition 3.1.12. *For a finite family $\{f_k\}_{k=1}^n$ of measurable functions with common domain E, the functions $\max\{f_1,\dots,f_n\}$ and $\min\{f_1,\dots,f_n\}$ are measurable.*

Proof. Let $f(x) = \max_{1\le i\le n} f_i(x)$, then $f(x) > c \iff$ there is $j \in \{1,2,\dots,n\}$ such that $f_j(x) > c$, hence

$$\{x \in E : f(x) > c\} = \bigcup_{j=1}^{n} \{x \in E : f_j(x) > c\}.$$

Similarly $\left\{x \in E : \min\limits_{1\le i\le n} f_i(x) < c\right\} = \bigcup\limits_{j=1}^{n} \left\{x \in E : f_j(x) < c\right\}.$ □

Any function f can be split by its **positive part** and the **negative part** defined by $f^+(x) = \max\{f(x),0\}$ and $f^-(x) = \max\{-f(x),0\}$ respectively. Here f^+ and f^- are both nonnegative and their difference is f. That is,

$$f = f^+ - f^-.$$

When f is measurable, Proposition 3.1.12 tells us f^+ and f^- are also measurable. The decomposition plays an important role in defining Lebesgue integral of measurable functions. Sometimes a proof can be simplified by noticing that properties possessed by nonnegative measurable functions may also be translated to general measurable ones by a careful modification.

3.2 Sequential Pointwise Limits

Let's define our common terminology as follows, these two modes of convergence are studied in mathematical analysis course.

Definition 3.2.1. Let $\{f_n\}$ be a sequence of functions with common domain E, a function f on E and a set $A \subseteq E$, we say that

(i) The sequence $\{f_n\}$ converges to f (denoted by $f_n \to f$) pointwise on A provided
$$\lim_{n\to\infty} f_n(x) = f(x), \quad \forall x \in A.$$

(ii) The sequence $\{f_n\}$ converges to f (denoted by $f_n \to f$) pointwise a.e. on A provided it converges to f pointwise on $A \setminus B$ with $m(B) = 0$.

(iii) The sequence $\{f_n\}$ converges to f uniformly (denoted by $f_n \rightrightarrows f$) on A provided for each $\epsilon > 0$, there is an N such that
$$n > N \implies |f - f_n| < \epsilon \text{ on } A,$$
where $|f - f_n| < \epsilon$ on A means $|f(x) - f_n(x)| < \epsilon$ for all $x \in A$.

In other courses we may run into other modes of convergence (some of which cannot be described by balls induced by metric), for example, norm convergence, weak convergence (both in L^p space) and also convergence in measure. To distinguish them they are denoted by $\xrightarrow{\|\cdot\|}$, $\xrightarrow{w}$ and $\xrightarrow{m}$ respectively. Usually by $f_n \to f$ (without the word "pointwise") we mean norm convergence when the collection of functions of interest is normed[(1)].

Proposition 3.2.2. *Let $\{f_n\}$ be a sequence of measurable functions on E such that $f_n \to f$ pointwise a.e. on E, then f is measurable.*

Proof. By possibly excising a set of measure zero from E, let's assume $f_n \to f$ pointwise on all of E. Let $c \in \mathbb{R}$, we see that $f(x) > c \iff \exists n \in \mathbb{N}, \exists N \in \mathbb{N}, \forall j \geq N, f_j(x) > c + \frac{1}{n}$, thus

$$\{x \in E : f(x) > c\} = \bigcup_{n=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{j=N}^{\infty} \left\{x \in E : f_j(x) > c + \frac{1}{n}\right\}. \qquad \square$$

Proposition 3.2.3. *For a sequence $\{f_n\}$ of measurable functions with common domain E, each of the following functions is measurable.*

$$\inf_{n\geq 1} f_n, \quad \sup_{n\geq 1} f_n, \quad \varliminf_{n\to\infty} f_n \quad \text{and} \quad \varlimsup_{n\to\infty} f_n.$$

Proof. Let $c \in \mathbb{R}$, $\inf_{n\geq 1} f_n(x) < c$ iff $f_j(x) < c$ for some $j \in \mathbb{N}$. We note also that $\varliminf_{n\to\infty} f_n(x) = \lim_{n\to\infty} \inf_{j\geq n} f_j(x)$ and $\varlimsup_{n\to\infty} f_n(x) = \lim_{n\to\infty} \sup_{j\geq n} f_j(x)$. The rest is left as exercises. $\square$

3.3 Simple Approximation

Definition 3.3.1. A real-valued function φ defined on a measurable set E is called **simple** provided it is measurable and takes only a finite number of values.

Definition 3.3.2. For any subset A of $\mathbb{R}$, the **characteristic function of** A, χ_A[(2)], is defined as follows:

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

It can be verified directly that χ_A is measurable if and only if A is measurable[(3)].

Assume φ is measurable and only takes the values $a_1, a_2, \ldots, a_n$ on E. By defining $A_i = \varphi^{-1}(a_i)$ for $i = 1, 2, \ldots, n$, we have $\bigsqcup_{i=1}^n A_i = E$ and a **canonical representation** of φ:

$$\varphi = \sum_{i=1}^n a_i \chi_{A_i}.$$

(1) We have seen an example in the first chapter that d_∞ is a norm on $C[a,b]$.

(2) Also denoted by $\mathbf{1}_A$ and also called indicator function.

(3) So we have our first example of nonmeasurable function, the characteristic function of a nonmeasurable set.

In particular, we call a simple function a step function when A_i are open intervals.

Definition 3.3.3. A **step function** on $[a,b]$ is a function s such that $s(x) = c_i$ for $x_{i-1} < x < x_i$ and the collection $\{x_0, x_1, \ldots, x_n\}$ forms a partition of $[a,b]$ ($x_0 = a, x_n = b$).

Lemma 3.3.4 (Simple Approximation). *Let f be a measurable real-valued function on E. Assume f is bounded on E, then for each $\epsilon > 0$, there are simple functions φ_ϵ and ψ_ϵ defined on E which have the following approximation properties:*

$$\varphi_\epsilon \le f \le \psi_\epsilon \quad \textit{and} \quad 0 \le \psi_\epsilon - \varphi_\epsilon < \epsilon \text{ on } E.$$

Proof. Let $\epsilon > 0$ be given and $[c,d)$ a bounded interval that contains $f(E)$. Let $\{y_0, y_1, \ldots, y_n\}$ partition $[c,d)$, where $c = y_0 < y_1 < y_2 < \cdots < f_n = d$ and $y_i - y_{i-1} < \epsilon$, for $i = 1,2,\ldots,n$. Define $A_i = f^{-1}[y_{i-1}, y_i)$, then $\{A_i\}_{i=1}^n$ is a disjoint collection of subsets of E such that $\bigsqcup_{i=1}^n A_i = E$. Now construct simple functions

$$\varphi_\epsilon = \sum_{i=1}^n y_{i-1}\chi_{A_i} \quad \text{and} \quad \psi_\epsilon = \sum_{i=1}^n y_i \chi_{A_i}.$$

Since for each $x \in E$, $x \in A_k$ for some k and thus $f(x) \in [y_{k-1}, y_k)$, this implies

$$\varphi_\epsilon(x) = y_{k-1} \le f(x) < y_k = \psi_\epsilon(x),$$

and $0 \le \psi_\epsilon - \varphi_\epsilon \le \max_{1 \le i \le n}(y_i - y_{i-1}) < \epsilon$. □

Remark. The simple approximation lemma tells us every bounded measurable function is a uniform limit of a sequence of simple functions. Also from the proof of the lemma when f is nonnegative, we can choose $\varphi_\epsilon \ge 0$.

Theorem 3.3.5 (Simple Approximation). *An extended real-valued function f on a measurable set E is measurable if and only if there is a sequence of simple functions $\{\varphi_n\}$ on E for which $\varphi_n \to f$ pointwise on E and has the property that*

$$|\varphi_n| \le |f| \textit{ on } E \textit{ for all } n.$$

If f is nonnegative, we may choose $\{\varphi_n\}$ to be increasing with $\varphi_n \ge 0$.

Proof. By Proposition 3.2.2 the if-direction is clear. For the converse let's first assume $f \ge 0$, the general case is left as exercise. Define $E_n = \{x \in E : f(x) \le n\}$ where $n \in \mathbb{N}$, then $f|_{E_n}$ is a bounded measurable function on E_n and thus by simple approximation lemma there are φ_n and ψ_n on E_n such that

$$0 \le \varphi_n \le f|_{E_n} \le \psi_n \quad \text{and} \quad 0 \le \psi_n - \varphi_n < \tfrac{1}{n}.$$

Extend φ_n on $E \setminus E_n$ by defining $\varphi_n|_{E \setminus E_n} \equiv n$, we claim that φ_n is our desired simple function.

Case 1. Let $x \in E$ be such that $f(x) = \infty$, then $x \notin E_n$ for all n, hence $\varphi_n(x) = n$, thus $\lim_{n\to\infty} \varphi_n(x) = \infty = f(x)$.

Case 2. Let $x \in E$ be such that $f(x) < \infty$, then $x \in E_N$ for some $N \in \mathbb{N}$, so for all $j \ge N$, $x \in E_j \implies 0 \le f(x) - \varphi_j(x) < \frac{1}{j}$, hence $\lim_{j\to\infty} \varphi_j(x) = f(x)$.

Now the general result follows from considering the positive and negative parts.

In case when f is nonnegative, we may replace φ_n by $\phi_n := \max\{\varphi_1,\varphi_2,\dots,\varphi_n\}$ to get an increasing sequence of simple functions. $\square$

Remark. As $\mathbb{R} = \bigcup_{n\in\mathbb{N}}[-n,n]$ and each $m([-n,n])$ is finite, thus $\mathbb{R}$ is σ-finite[(4)] and hence we can replace each φ_n by $\phi_n := \varphi_n\chi_{[-n,n]}$. That is to say, we can further assume each φ_n vanishes outside a set of finite measure. We describe such functions have finite support.

3.4 Littlewood's Three Principles

For Lebesgue measure Littlewood's Three Principles are roughly the following.

- Every (measurable) set is "nearly" a finite union of open intervals (Theorem 3.4.1);
- Every pointwise convergent sequence of (measurable) functions is "nearly" uniformly convergent (Egoroffs Theorem);
- Every (measurable) function is "nearly" continuous (Lusin's Theorem).

It is worth noting that among the three principles Egoroff's Theorem can be generalized to arbitrary finite measure space (X,Σ,μ)[(5)].

Theorem 3.4.1 (The First Principle). *Let E have finite measure. Then for each $\epsilon > 0$, there is a finite disjoint collection of bounded open intervals $\{I_k\}_{k=1}^n$ such that*

$$m(E\Delta(\textstyle\bigsqcup_{k=1}^n I_k)) < \epsilon.$$

Proof. Let $\epsilon > 0$ be given. Since $m^*(E) = m(E) < \infty$, there is an open $U \supseteq E$ such that $\lambda(U) - m(E) < \frac{\epsilon}{2}$. Write $U = \bigsqcup(a_i,b_i)$, if the union is already finite, done. Assume the union is infinite, then there is an N such that $\bigsqcup_{i=N+1}^\infty \lambda((a_i,b_i)) < \epsilon/2$, write $O = \bigsqcup_{i=1}^N(a_i,b_i)$, now

$$m(O\setminus E) \le m(U\setminus E) < \epsilon/2$$

and

$$m(E\setminus O) \le m(U\setminus O) = \sum_{i=N+1}^{\infty} \lambda((a_i,b_i)) < \epsilon/2. \qquad \square$$

Theorem 3.4.2 (Egoroff). *Assume E has finite measure. Let $\{f_n\}$ be a sequence of measurable functions on E that converges pointwise on E to the real-valued function f. Then for each $\epsilon > 0$, there is a closed set F contained in E for which*

$$f_n \rightrightarrows f \text{ on } F \quad \text{and} \quad m(E\setminus F) < \epsilon.$$

[(4)]A measure space (X,Σ,μ) is said to be σ-finite if there are $X_1, X_2, \dots \in \Sigma$ such that $X = \bigcup_{i=1}^\infty X_i$ and $\mu(X_i) < \infty$ for each i.

[(5)]One of the troubles in translating facts in terms of Lebesgue measure to general measure space X is that X may not be complete. That is, the σ-algebra of subsets of X does not necessarily contain all subsets of a set of measure zero. In such incomplete measure space we do not have $f = g$ a.e. on X and f measurable $\implies g$ measurable (Proposition 3.1.8), nor $\{f_n\}$ a sequence of measurable functions and $\lim_{n\to\infty} f_n = f$ a.e. on $X \implies f$ measurable (Proposition 3.2.2).

Proof. This theorem actually holds in a much general setting, we defer the proof to Theorem 5.1.13 with a little change of notations. $\square$

Let $A \subseteq \mathbb{R}$ be a subset and f a real-valued function defined on A. The upcoming proposition requires us to recall what is meant by continuity on arbitrary A, where A is given so called subspace (metric) topology (the collection of "open" sets induced by that in the larger space containing it). Continuity between metric spaces is described simply by balls which is still what we concern the most in the subspace[6]. Those open balls in the subspace A are of the form

$$\begin{aligned} B_A(a,r) &:= \{x \in A : d(x,a) < r\} \\ &= \{x \in \mathbb{R} : d(x,a) < r\} \cap A \\ &= B(a,r) \cap A. \end{aligned}$$

Moreover, given metric spaces X and Y, a function $f : X \to Y$ is said to be continuous if and only if for each $x \in X$ and for any open ball V in Y containing $f(x)$, there is an open ball U in X containing x such that

$$f(U) \subseteq V. \tag{3.4.3}$$

In case when it holds at a point $x \in X$, we say that the function f is continuous at x.

This criterion is also true when U and V are topological bases elements. Metric space always has a metric topology (a natural topology generated by the collection of balls, an example of base). The similar criterion holds for general topological spaces (where V is open set in Y and U is open set in X). See page 104 of the book TOPOLOGY (2^{nd} edition) written by James R. Munkres.

Remark. For metric spaces X,Y, a continuous $f : X \to Y$ and a $Z \subset X$, the restriction $f|_Z : Z \to Y$ is also continuous (of course, with respect to subspace topology of Z). This is easily proved by the ball-ball argument given in (3.4.3). This is also true for general topological spaces. Note that "f is continuous on Z" and "$f|_Z$ is continuous" are totally different matters.

Lemma 3.4.4. *Let f be a simple function defined on E. Then for each $\epsilon > 0$, there is a continuous functions g on $\mathbb{R}$ and a closed set F contained in E for which*

$$f = g \text{ on } F \quad \text{and} \quad m(E \setminus F) < \epsilon.$$

Proof. Let $f = \sum_{i=1}^n a_i \chi_{A_i}$, where $\bigsqcup_{i=1}^n A_i = E$. Then for each i there is a closed $F_i \subseteq A_i$ such that $m(A_i \setminus F_i) < \frac{\epsilon}{n}$, it implies $m(E \setminus F) < \epsilon$, where $F = \bigsqcup F_i$ (clearly closed). It remains to define a continuous function on $\mathbb{R}$ that agrees with f on F.

Let's define $g := \sum_{i=1}^n a_i \chi_{F_i}$, then $g = f$ on F. We now try to prove that $g|_F$ is continuous on F, after that by Problem 3.13 we can extend g to a continuous function on $\mathbb{R}$. For each $x \in F$ there is an i such that $x \in F_i$, but this implies $x \in \mathbb{R} \setminus \bigsqcup_{j \neq i} F_j$, meaning there is a $\delta > 0$ such that $B(x,\delta) \subseteq \mathbb{R} \setminus \bigsqcup_{j \neq i} F_j$, then $B(x,\delta) \cap F \subseteq F_i$ on which g is constant and thus by the criterion in (3.4.3), $g|_F$ is continuous on F. $\square$

Theorem 3.4.5 (Lusin). *Let f be a real-valued measurable function on E. Then for each $\epsilon > 0$, there is a continuous function g on $\mathbb{R}$ and a closed $F \subseteq E$ for which*

$$f = g \text{ on } F \quad \text{and} \quad m(E \setminus F) < \epsilon.$$

[6] Recall that a subspace of a metric space is still a metric space.

Proof. We only prove the case when $m(E) < \infty$, the case that $m(E) = \infty$ is left as exercise.

As f is measurable on E, by the simple approximation theorem there is a sequence of simple functions $\{\phi_n\}$ such that $\phi_n \to f$ pointwise on E. Now by Lemma 3.4.4 for each $n \in \mathbb{N}$ there is a closed $F_n \subseteq E$ and a continuous function g_n on $\mathbb{R}$ such that

$$\phi_n = g_n \text{ on } F_n \quad \text{and} \quad m(E \setminus F_n) < \frac{\epsilon}{2^{n+1}}.$$

Let $F_0 = \bigcap_{n=1}^\infty F_n$, then $\phi_n \to f$ pointwise on F_0, $\phi_n = g_n$ on F_0 and $m(E \setminus F_0) < \epsilon/2$. In order for f be equal to a continuous function pointwise on some subset, we expect the convergence ϕ_n to be uniform. Here Egoroff's theorem can do the task.

By Egoroff's theorem, there is a closed $F \subseteq F_0$ such that $\phi_n \rightrightarrows f$ on F and $m(F_0 \setminus F) < \epsilon/2$. Moreover, the inequality

$$\begin{aligned}|f(x) - f(y)| &\le |f(x) - \phi_n(x)| + |\phi_n(x) - \phi_n(y)| + |\phi_n(y) - f(y)| \\ &= |f(x) - \phi_n(x)| + |g_n(x) - g_n(y)| + |\phi_n(y) - f(y)|\end{aligned}$$

implies $f|_F$ is continuous on F (detail can be found in the following remark), of course, with respect to the subspace topology, so by Problem 3.13 we can extend $f|_F$ to a continuous function on $\mathbb{R}$. Finally $m(E \setminus F) = [m(E) - m(F_0)] + [m(F_0) - m(F)] < \epsilon$. $\square$

Remark.

- The continuity of $f|_F$ on F can be argued as follows.

  Let $x_0 \in F$ be fixed, then for each $\epsilon > 0$, the uniform convergence of $\{\phi_n\}$ on F implies there is an N such that $|f - \phi_N| < \epsilon/3$ on F. For this choice of N there is a $\delta > 0$ such that $y \in B(x_0,\delta) \implies |g_N(x_0) - g_N(y)| < \epsilon/3$. Hence when $y \in B(x_0,\delta) \cap F$, $|f(x_0) - f(y)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$. The implication actually means for any $\epsilon > 0$, there is a $\delta > 0$ such that

  $$f|_F(B(x_0,\delta) \cap F) \subseteq B(f(x_0),\epsilon).$$

  That satisfies the criterion given in (3.4.3).

- Sometimes when $m(A \setminus B), m(B \setminus C)$ are small, we expect $m(A \setminus C)$ is also small. The approach used in the last line of the proof in Lusin's theorem is not applicable in general as it requires $m(F_0)$ and $m(F)$ be finite. To overcome this difficulty we observe that for any set A, B and C, $A \setminus C \subseteq (A \setminus B) \sqcup (B \setminus C)$!

- To prove Lusin's theorem in the case that $m(E) = \infty$ it is not a good way to extend the result on $E_n := E \cap [-n,n]$ successively.

- By the method we use in Problem 3.13 we see that for closed L and a continuous function $f : L \to \mathbb{R}$, the extension $F : \mathbb{R} \to \mathbb{R}$ of f can be chosen so that

  $$|F(x)| \le \sup |f(L)|.$$

  In other words, when f is bounded measurable and real-valued on E, there is a continuous extension F on $\mathbb{R}$ whose magnitude is as large as f.

As an application of Lusin's theorem let's identify all real-valued measurable functions on $\mathbb{R}$ that preserves addition.

Proposition 3.4.6. *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function such that for any $x, y \in \mathbb{R}$,*

$$f(x+y) = f(x) + f(y),$$

then f is a continuous function on $\mathbb{R}$.

It is a simple exercise in mathematical analysis that such continuous functions must be of the form $f(x) = f(1)x$. We thereby complete the classification of additive measurable functions.

Proof. Let's first observe that $f(0) = 0$ and $f(x+h) - f(x) = f(h)$, it follows that f is continuous on $\mathbb{R}$ iff f is continuous at 0, let's focus on neighborhoods of 0. By Lusin's theorem, there is a measurable $E \subseteq \mathbb{R}$ with $m(E) > 0$ such that there is $g \in C(\mathbb{R})$ and $f|_E = g|_E$. We can find a compact $K \subseteq E$ such that $m(K) > 0$, then since g is uniformly continuous on K, for any $\epsilon > 0$, we can find $\delta > 0$ so that

$$x, y \in K, |x-y| < \delta \implies |f(x-y)| = |f(x) - f(y)| < \epsilon.$$

We are almost done, by Steinhaus Theorem 2.10.9, 0 is an interior point of $K - K$, hence there is $\delta' > 0$ such that $(-\delta', \delta') \subseteq K - K$. So whenever $|z| < \min\{\delta, \delta'\}$, there are $x, y \in K$ such that $|z| = |x-y| < \delta$, and

$$|f(z)| = |f(x-y)| < \epsilon,$$

from which we conclude f is continuous at 0. $\square$

3.5 Exercises and Problems

Exercises

3.1. Prove that: For a sequence $\{f_n\}$ of measurable functions with common domain E, each of the following functions is measurable.

$$\inf_{n\geq 1} f_n, \quad \sup_{n\geq 1} f_n, \quad \underline{\lim}_{n\geq 1} f_n \quad \text{and} \quad \overline{\lim}_{n\geq 1} f_n.$$

3.2. Suppose f is a real-valued function on $\mathbb{R}$ such that $f^{-1}(c)$ is measurable for each $c \in \mathbb{R}$, is f necessarily measurable?

3.3. Prove that if f is measurable on E, then so is f^2. Let g be defined on a measurable E and g^2 be measurable, show that if $\{x \in E : g(x) > 0\}$ is measurable, then g is measurable on E.

3.4. Prove the following:

(a) Let $\mathcal{F}$ be a family of continuous functions on $(0,1)$, show that

$$g(x) = \sup\{f(x) : f \in \mathcal{F}\} \quad \text{and} \quad h(x) = \inf\{f(x) : f \in \mathcal{F}\}$$

are measurable functions on $(0,1)$.

(b) For every $n \in \mathbb{N}$, let $f_n : \mathbb{R} \to [0,1]$ be a measurable function, show that the set

$$A = \left\{ x \in \mathbb{R} : \lim_{n\to\infty} \frac{f_1(x) + 2f_2(x) + \cdots + nf_n(x)}{n^2} \text{ does not exist} \right\}$$

is measurable.

3.5. Prove the following:

(a) From Section 2.10.2 we know that there is a strictly increasing and continuous $h : [0,1] \to [0,2]$ with $h(0) = 0$ and $h(1) = 2$ which maps a set $N \subseteq$ Cantor set $\subseteq [0,1]$ onto a nonmeasurable set $h(N)$. We extend h to $H : \mathbb{R} \to \mathbb{R}$ which is strictly increasing and maps $\mathbb{R}$ onto $\mathbb{R}$ (let's say we extend it linearly with constant positive slope), then we get a continuous inverse H^{-1}. Construct $f : \mathbb{R} \to \{0,1\}$ as follows (recall that $H(N) = h(N)$)

$$f(x) = \chi_N \circ H^{-1}(x),$$

show that f being a composition of measurable functions is not measurable.

(b) Suppose g and h are real-valued functions defined on all of $\mathbb{R}$, g is measurable and h is continuous. Is the composition $g \circ h$ necessarily measurable?

3.6. Let I be an interval and $f : I \to \mathbb{R}$ be increasing. Show that f is measurable by first showing that, for each natural number n, the strictly increasing function $x \mapsto f(x) + \frac{x}{n}$ is measurable.

3.7. Let g be a mapping from $\mathbb{R}$ onto $\mathbb{R}$ for which there is a constant $c > 0$ such that

$$|g(u) - g(v)| \geq c|u - v|, \quad \forall u, v \in \mathbb{R}.$$

Show that if $f : \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable, then so is the composition $f \circ g : \mathbb{R} \to \mathbb{R}$.

3.8. Prove the following:

(a) The product and linear combination of finitely many simple functions on E is still a simple function.

(b) The product and linear combination of finitely many step functions on an interval I is still a step function.

3.9. Prove the following approximation properties:

(a) Let I be a compact interval and E a measurable subset of I. Let $\epsilon > 0$, show that there is a step function h on I and a measurable subset F of I for which

$$h = \chi_E \text{ on } F \quad \text{and} \quad m(I \setminus F) < \epsilon.$$

[**Hint:** Use the first principle.]

(b) Let I be a compact interval and ψ a simple function defined on I. Let $\epsilon > 0$. Show that there is a step function h on I and a measurable subset F of I for which

$$h = \psi \text{ on } F \quad \text{and} \quad m(I \setminus F) < \epsilon.$$

If $m \le \psi \le M$, then we can take h so that $m \le h \le M$. That is to say, each simple function on E is "nearly" a step function.

(c) Let I be a compact interval and f a bounded measurable function defined on I. Let $\epsilon > 0$. Show that there is a step function h on I and a measurable subset F of I for which

$$|f - h| < \epsilon \quad \text{and} \quad m(I \setminus F) < \epsilon.$$

[**Hint:** Recall that step function φ on $[a,b]$ has a canonical representation $\varphi = \sum_{i=1}^{n} a_i \chi_{I_i}$, where I_i are bounded interval.]

3.10. Let E have finite measure and f be a measurable function on E that is finite a.e.. Prove that given $\epsilon > 0$, there is a subset F of E such that

$$f \text{ is bounded on } F \quad \text{and} \quad m(E \setminus F) < \epsilon.$$

That is to say, each measurable function that is finite a.e. on a set of finite measure is "nearly" a bounded measurable function.

Problems

3.11. Express a measurable function as the difference of nonnegative measurable functions and thereby prove the general simple approximation theorem based on the special case of nonnegative measurable function.

3.12. Show that the conclusion of Egoroff's Theorem can fail if we drop the assumption that the domain has finite measure.

3.13. Suppose f is a function that is continuous on a closed subset F of $\mathbb{R}$. Show that f has a continuous extension to all of $\mathbb{R}$ (this is a special case of the Tietze Extension Theorem 7.1.4).
[**Hint:** Express $\mathbb{R} \setminus F$ as the union of a countable disjoint collection of open intervals and define f to be linear on the closure of each of these intervals.]

3.14. Prove the extension of Lusin's Theorem to the case that E has infinite measure.

3.15. Let $\{f_n\}$ be a sequence of measurable functions on $[a,b]$ and f a real-valued function on $[a,b]$ such that

(i) $|f_n| \le M_n$ for $n = 1, 2, \dots$

(ii) $\lim_{n\to\infty} f_n = f$ a.e. on $[a,b]$.

Show that given any $\delta > 0$, there is a measurable $E \subseteq [a,b]$ and a constant M such that $m(E) < \delta$ and $|f|, |f_1|, |f_2|, \cdots \le M$ on $[a,b] \setminus E$.

Part II

Summer 2011-2012

Chapter 4

General Measure

In summer 2010-2011 we discussed Lebesgue measure m on $\mathbb{R}$, m-measurable function and the Littlewood's Three Principles on $\mathbb{R}$. We also discussed some "geometric structure" of measurable set on $\mathbb{R}$ with positive measure. However, we haven't discussed integration and differentiation theory on $\mathbb{R}$.

This year we aim to discuss general measure. After broadening our view point on measure, we try to go back to measure theory on $\mathbb{R}^n$.

Since we seldom deal with the sets like $\{a-b : a \in A, b \in B\}$ (which we have encountered in measure theory on $\mathbb{R}$). From now on the notation "$-$" between sets is reserved for set complement, i.e., for two sets A and B, $A - B := A \setminus B = \{x \in A : x \notin B\}$.

4.1 Outer Measure

In learning Lebesgue measure we start with defining "length" on intervals, we next define outer and inner measures in terms of "length". After that we have shown that Lebesgue measurability is independent of inner measure. This suggests us a general theory can be built by outer measure alone.

We now define outer measure and next define measurability of subsets with respect to such measure.

Definition 4.1.1. An **outer measure** on a set X is a set function $\mu^* : 2^X \to [0,\infty]$ such that:

(i) $\mu^*(\emptyset) = 0.$ (Nonnegativeness)

(ii) $A \subseteq B \implies \mu^*(A) \le \mu^*(B).$ (Monotonicity)

(iii) $\mu^*(\bigcup_{i=1}^\infty A_i) \le \sum_{i=1}^\infty \mu^*(A_i).$ (Subadditivity)

At the moment we don't try to construct examples of outer measures. As we shall see shortly in Theorem 4.3.1, there are abundant examples of outer measures generated by set functions defined on any collection of subsets. Concrete examples will be constructed after that. Let's first investigate basic properties of outer measures.

Definition 4.1.2. Given an outer measure μ^* on X, a subset A of X is said to be **μ^*-measurable** provided for every subset Y of X,

$$\mu^*(Y) = \mu^*(Y \cap A) + \mu^*(Y - A).$$

In other words, A is μ^*-measurable iff A and $X - A$ can be used to split the outer measure of any subset of X. As an immediate consequence of the definition, a subset A of X is μ^*-measurable iff $X - A$ is μ^*-measurable.

All properties possessed by Lebesgue outer measure can be immediately translated to abstract outer measures, so are their proofs (we just need to replace the letter m by μ). We state the results here without repeating the proofs.

Proposition 4.1.3. *Let μ^* be an outer measurable defined on X.*

(i) The union of a finite collection of μ^*-measurable sets is μ^*-measurable.

(ii) The countable union of μ^*-measurable sets is μ^*-measurable.

(iii) The countable intersection of μ^*-measurable sets is μ^*-measurable.

*(iv) **(Countable Additivity)** Let $\{A_k\}_{k=1}^\infty$ be a disjoint collection of μ^*-measurable sets, then*

$$\mu^*\left(\bigsqcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu^*(A_k).$$

We have define σ-algebra in Definition 2.10.1, here we give another equivalent formulation:

Definition 4.1.4. Let X be a set. A ***σ*-algebra** on X is a collection $\Sigma \subseteq 2^X$ satisfying the following properties:

(i) $X \in \Sigma$.

(ii) $A, B \in \Sigma \implies A - B \in \Sigma$.

(iii) $A_1, A_2, \cdots \in \Sigma \implies \bigcup_{i=1}^\infty A_i \in \Sigma$.

Part (i) and (iii) are same as before, given these two, (ii) in Definition 2.10.1 and (ii) in Definition 4.1.4 are equivalent, i.e., $X - A \in \Sigma$ for all $A \in \Sigma$ iff $A - B \in \Sigma$ for all $A, B \in \Sigma$.

By definition of μ^*-measurability and Proposition 4.1.3 the collection of μ^*-measurable subsets of X, Σ, forms a σ-algebra. $\mu := \mu^*|_\Sigma$ is called the **measure induced by μ^***. So a "measure" can be defined once we have an outer measure.

Proposition 4.1.5. *Let μ^* be an outer measure on X, μ a measure induced by μ^*, then μ satisfies the following properties:*

(i) If $\mu^*(A) = 0$, then A is μ^*-measurable. Moreover, any $B \subseteq A$ is also μ^*-measurable with $\mu(B) = 0$.

(ii) $\mu(A) \geq 0$ if A is μ^*-measurable.

(iii) $\mu(\bigsqcup A_i) = \sum \mu(A_i)$ if A_i's are μ^*-measurable and disjoint.

Proof. For any subset Y of X, $\mu^*(Y) \leq \mu^*(Y \cap B) + \mu^*(Y - B) = \mu^*(Y - B) \leq \mu^*(Y)$, so (i) follows. (ii) follows from the definition of outer measure. (iii) follows from (iv) of Proposition 4.1.3. □

We come back to the discussion of outer measure after we have some notion of measure. We will see that outer measure is a main tool to extend some "special set function" (called premeasure) defined on some "special collection" (called semiring), to a measure defined on a σ-algebra containing this "special collection".

4.2 Measure, Measure Spaces and Their Completion

Definition 4.2.1. Let X be a space and let Σ be a σ-algebra of subsets of X. The couple (X,Σ) is called a **measurable space**.

We expect every measure should have properties (ii) and (iii) in Proposition 4.1.5, let's extract them as our definition of measure.

Definition 4.2.2. A **measure** on a measurable space (X,Σ) is a set function $\mu : \Sigma \to [0,\infty]$ satisfying the following properties:

(i) $\mu(\emptyset) = 0$.

(ii) $\mu(A) \geq 0$ for all $A \in \Sigma$.

(iii) $\mu(\bigsqcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ for pairwise disjoint $A_i \in \Sigma$.

By **measure space** we mean the triple (X,Σ,μ). That is, a measurable space together with a measure. A set $A \subseteq X$ is said to be **measurable** if $X \in \Sigma$

Example 4.2.3. The measure induced by μ^* is within a class of measure. □

Example 4.2.4 (Counting Measure). Let X be nonemtpy and consider the measurable space $(X,2^X)$. For each $A \in 2^X$, define

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

μ is a measure on X. To prove countable additivity, consider the collection of sets $\{A_k\}_{k=1}^{\infty}$ in X. If infinitely many of A_k's are nonempty, then both sides of $\mu(\bigsqcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k)$ is ∞. Suppose only finitely many of A_k's are nonempty, say $A_k = \emptyset$ for $k > n$. Then the equality $\mu(\bigsqcup_{k=1}^{n} A_k) = \sum_{k=1}^{n} \mu(A_k)$ holds obviously. □

Example 4.2.5 (Dirac Measure/Point Mass). Let X be nonempty and consider any σ-algebra, Σ, on X and consider (X,Σ). Fix an $x \in X$, for each $A \in \Sigma$ define

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

The countable additivity can be easily verified. μ_x is called the Dirac measure concentrated at x. Furthermore, for $a_1,a_2,\cdots \geq 0$ and $x_1,x_2,\cdots \in X$ the set function defined by

$$\delta(x) = \sum_{k=1}^{\infty} a_k \delta_{x_k}(x)$$

is called a **discrete measure**. The countable additivity follows from rearrangement of nonnegative series. Moreover, we also say that δ places a point mass a_k at x_k. □

Proposition 4.2.6. *Let (X,Σ,μ) be a measure space. Then μ has the following properties:*

(i) Finitely Additive: $\mu(\bigsqcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$.

(ii) Monotone: If $A \subseteq B$*, then* $\mu(A) \le \mu(B)$.

(iii) Excision Property: If $A \subseteq B, \mu(A) < \infty$*, then* $\mu(B - A) = \mu(B) - \mu(A)$.

(iv) Subadditive: $\mu(\bigcup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} \mu(A_i)$.

Remark. Monotone property (ii) + subadditivity (iv) is equivalent to **countable monotonicity**: If $A, B_i \in \Sigma$ and $A \subseteq \bigcup_{i=1}^{\infty} B_i$, then $\mu(A) \le \sum_{i=1}^{\infty} \mu(B_i)$.

Proof. (i) is true by taking $A_k = \emptyset$ for $k > n$ in the definition of countable additivity. (ii) is true since $\mu(B) = \mu(A) + \mu(B - A) \ge \mu(A)$, and (iii) is true due to the same identity.

Finally, let $B_1 = A_1$ and $B_i = A_i - \bigcup_{k=1}^{i-1} A_k$ for $i \ge 2$. Then $\bigcup B_i = \bigcup A_i$ and $\{B_i\}$ is a disjoint collection, hence $\mu(\bigcup A_i) = \mu(\bigcup B_i) = \sum \mu(B_i) \le \sum \mu(A_i)$, so (iv) is true. ☐

Proposition 4.2.7 (Continuity of Measure). *Let* (X, Σ, μ) *be a measure space.*

(i) If $\{A_k\}$ *is an ascending collection of measurable sets, then*

$$\mu\Big(\lim_{k\to\infty} A_k\Big) := \mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_{k\to\infty} \mu(A_k).$$

(ii) If $\{A_k\}$ *is a descending collection of measurable sets and* $\mu(A_k) < \infty$*, for some* k*, then*

$$\mu\Big(\lim_{k\to\infty} A_k\Big) := \mu\left(\bigcap_{k=1}^{\infty} A_k\right) = \lim_{k\to\infty} \mu(A_k).$$

Proof. We may copy the proof of Theorem 2.7.3, i.e., replace m^* by μ. The key property is countable additivity. ☐

Lemma 4.2.8 (Borel-Cantelli). *Let* (X, Σ, μ) *be a measure space. Let* $\{A_k\}_{k=1}^{\infty}$ *be a countable collection of measurable sets for which* $\sum_{k=1}^{\infty} \mu(A_k) < \infty$*. Then almost all* $x \in X$ *belong to at most finitely many of the* A_k*'s.*

Proof. The proof is same as the case that $\mu = m$. ☐

We mention a useful condition on a measure space, on which many results on finite measure space can be extended.

Definition 4.2.9. Let (X, Σ, μ) be a measure space. A subset A of X is **σ-finite** if it is contained in a countable union of sets of finite measure. The measure μ is **σ-finite** if the whole space X is σ-finite.

If the measure μ is σ-finite, we say that (X, Σ, μ) is a **σ-finite measure space**. In case $\mu(X) < \infty$, (X, Σ, μ) is called a **finite measure space**.

Example 4.2.10. Let $\mathcal{L}$ denote the collection of Lebesgue measurable subsets of $\mathbb{R}$, then $(\mathbb{R}, \mathcal{L}, m)$ is a σ-finite measure space, $\mathbb{R}$ is σ-finite and m is a σ-finite measure on $\mathbb{R}$. ☐

Definition 4.2.11. Let (X,Σ,μ) be a measure space. The measure μ is **complete** if for every $A \in \Sigma$ with $\mu(A) = 0$, $B \subseteq A \implies B \in \Sigma$.

A measurable space equipped with a complete measure is called a **complete measure space**.

Example 4.2.12. The counting measure defined in Example 4.2.4 is always a complete measure. □

Example 4.2.13. By (i) of Proposition 4.1.5, a measure induced by an outer measure is always complete. Hence Lebesgue measure m is a complete measure since it is induced by Lebesgue outer measure m^*. □

Example 4.2.14. A simple incomplete measure space can be constructed by Dirac measure defined in Example 4.2.5. Consider a σ-algebra

$$\Sigma := \{\emptyset,\{1\},\{2,3\},\{1,2,3\}\}$$

on $\{1,2,3\}$. Let $\mu_1(A) := \chi_A(1)$, then $\mu_1(\{2,3\}) = 0$ but $\{2\},\{3\} \notin \Sigma$. □

Example 4.2.15. Let $\mathcal{B}$ be the Borel σ-algebra on $\mathbb{R}$ and define $\mu = m|_{\mathcal{B}}$, then the measure space $(\mathbb{R},\mathcal{B},\mu)$ is not complete. To see this, recall that Cantor set C is Borel since it is closed and $\mu(C) = m(C) = 0$, but by Theorem 2.10.7 there is a subset of C that fails to be Borel. □

Each measure space can be completed by enlarging the existing σ-algebra. The way to achieve this is very natural. Suppose $\mu(Z) = 0$ and $B \subset Z$, we wish B was measurable, once this is true, for any $A \in \Sigma$, $A \cup B$ is necessarily measurable.

Proposition 4.2.16 (Completion of a σ-Algebra). *Let (X,Σ,μ) be a measure space. Let $\mathcal{Z} = \bigcup_{Z\in\Sigma,\mu(Z)=0} 2^Z$, define*

$$\overline{\Sigma} = \{A \cup B : A \in \Sigma, B \in \mathcal{Z}\}.$$

For $E \in \overline{\Sigma}$, i.e., $E = A \cup B$, for some $A \in \Sigma, B \in \mathcal{Z}$, we define $\overline{\mu}(E) = \mu(A)$. Then $\overline{\Sigma}$ is a σ-algebra containing Σ, $\overline{\mu}$ is a well-defined measure that extends μ, and $(X,\overline{\Sigma},\overline{\mu})$ is a complete measure space.

Before we begin the proof, let's fix the following choice: Let $A_i \cup B_i \in \overline{\Sigma}$, with $A_i \in \Sigma$ and $B_i \subseteq Z_i$, for some $Z_i \in \Sigma$ with $\mu(Z_i) = 0$, $i = 1,2,\ldots$. Note that

$$B \in \mathcal{Z} \iff B \subseteq Z, \text{ for some } Z \in \Sigma \text{ with } \mu(Z) = 0,$$

they are subsets that are "almost" measure zero.

Proof. We first show that $\overline{\Sigma}$ is a σ-algebra. First of all, $X \in \overline{\Sigma}$. Secondly,

$$\begin{aligned} &A_1 \cup B_1 - A_2 \cup B_2 \\ &= (A_1 - A_2 - Z_2) \cup \big((A_1 - A_2) \cap (Z_2 - B_2)\big) \cup (B_1 - A_2 - B_2), \end{aligned} \tag{4.2.17}$$

this shows that $A_1 \cup B_1 - A_2 \cup B_2 \in \overline{\Sigma}$, so $\overline{\Sigma}$ is closed under relative complement. Finally, $\bigcup(A_i \cup B_i) = (\bigcup A_i) \cup (\bigcup B_i) \in \overline{\Sigma}$. We conclude that $\overline{\Sigma}$ is a σ-algebra. Of course $\overline{\Sigma} \supseteq \Sigma$.

Next we show $\overline{\mu}$ is well-defined measure on $\overline{\Sigma}$, let $A_1 \cup B_1 = A_2 \cup B_2$, then by the set equality (4.2.17) each term in the union must be empty. In particular, $A_1 - A_2 - Z_2 = \emptyset$, hence $A_1 - A_2 \subseteq (A_1 - A_2) \cup Z_2 = (A_1 - A_2 - Z_2) \cup Z_2 = Z_2$. Switching 1 and 2, we have $A_2 - A_1 \subseteq Z_1$, hence

$$(A_1 - A_2) \sqcup (A_2 - A_1) \subseteq Z_1 \cup Z_2,$$

so $\mu(A_1 \Delta A_2) = 0$, which implies $\mu(A_1) = \mu(A_2)$, thus $\overline{\mu}(A_1 \cup B_1) = \overline{\mu}(A_2 \cup B_2)$. $\overline{\mu}$ extends μ by the equality $A = A \cup \emptyset$ for $A \in \Sigma$. The countable additivity of $\overline{\mu}$ inherits from that of μ.

It remains to show $\overline{\Sigma}$ is complete. Assume that $\overline{\mu}(A_1 \cup B_1) = 0$ and $E \subseteq A_1 \cup B_1$. By definition, $\mu(A_1) = 0$, and hence $E = \emptyset \cup E$ where $E \subseteq A_1 \cup Z_1$, with $\mu(A_1 \cup Z_1) = 0$, so $E \in \overline{\Sigma}$. □

Proposition 4.2.16 actually says that a completion can be obtained by inserting all "almost" measure zero subsets into Σ.

Definition 4.2.18. The measure space $(X, \overline{\Sigma}, \overline{\mu})$ defined in Proposition 4.2.16 is called the **completion** of the measure space (X, Σ, μ).

The completion is minimal in the following sense, which we leave the proof as an exercise for practice.

Proposition 4.2.19. *Let (X, Σ', μ') be another complete measure space that extends (X, Σ, μ) in the sense that $\Sigma' \supseteq \Sigma$ and $\mu'|_\Sigma = \mu$, and let $(X, \overline{\Sigma}, \overline{\mu})$ be the completion of (X, Σ, μ), then (X, Σ', μ') also extends $(X, \overline{\Sigma}, \overline{\mu})$.*

4.3 Extension to Measure

4.3.1 Construction of Outer Measure

In the past, we define Lebesgue outer measure of $E \subseteq \mathbb{R}$ to be the infimum of the lengths of open sets containing E. This is not an appropriate choice for general measure theory since measure space, as we have seen, needs not be an topological space.

However, each open sets O in $\mathbb{R}$ can be written as a union of open intervals, namely, $O = \bigsqcup I_i$, on which the "length" is easily assigned. Moreover, $\lambda(O) = \sum \lambda(I_i)$, we have

$$m^*(E) = \inf \left\{ \sum \lambda(I_i) : \bigcup I_i \supseteq E, I_i \text{ are open intervals} \right\},$$

here $\Sigma, \cup$ means countable summation and union. Theorem 4.3.1 states that this method of construction of outer measure on $\mathbb{R}$ works in general.

Theorem 4.3.1. *Let $\mathcal{S}$ be a collection of subsets of X such that $\emptyset \in \mathcal{S}$ and $\lambda : \mathcal{S} \to [0, \infty]$ a set function. Define $\lambda(\emptyset) = 0$ and in case A can be contained in a countable union of subsets in $\mathcal{S}$, define*

$$\lambda^*(A) = \inf \left\{ \sum \lambda(S_i) : \bigcup S_i \supseteq A, S_i \in \mathcal{S} \right\},$$

(here Σ and $\cup$ are countable) otherwise we define $\lambda^*(A) = \infty$. Then λ^* is an outer measure (induced by λ).

Proof. Since $\lambda(\emptyset) = 0$, $0 \le \lambda^*(\emptyset) \le \lambda(\emptyset) = 0$. To show λ^* is monotone, let $A \subseteq B$. If $\lambda^*(B) = \infty$, done. Otherwise let $\bigcup S_i \supseteq B$, for some $S_i \in \mathcal{S}$, then $\bigcup S_i \supseteq A$, hence $\lambda^*(A) \le \sum \lambda(S_i)$ for all cover $\{S_i\}$ of B. By taking infimum over all possible covers of B, one has $\lambda^*(A) \le \lambda^*(B)$.

Finally we need to show subadditivity. Let $\{A_i\}_{i\ge 1}$ be a collection of subsets in X. We need to show $\lambda^*(\bigcup A_i) \le \sum \lambda^*(A_i)$. If there is $\lambda^*(A_i) = \infty$, we are done. Assume $\lambda^*(A_i) < \infty$ for all i. Let $\epsilon > 0$ be given. then for each i we can choose a cover $\bigcup_j S_{ij} \supseteq A_i$ such that $\sum_j \lambda(S_{ij}) - \lambda^*(A_i) < \epsilon/2^i$. Since $\bigcup_i \bigcup_j S_{ij} \supseteq \bigcup A_i$, we have

$$\lambda^*\left(\bigcup_{i=1}^{\infty} A_i\right) \le \sum_{(i,j)\in\mathbb{N}\times\mathbb{N}} \lambda(S_{ij}) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \lambda(S_{ij}) \le \sum_{i=1}^{\infty} \lambda^*(A_i) + \epsilon.$$

Since the choice of ϵ can be relaxed, we are done. □

Also since Theorem 4.3.1 is a basis of this chapter, from now on $\mathcal{S}$ **always contains** $\{\emptyset\}$, unless otherwise specified.

Definition 4.3.2. In the same setting of Theorem 4.3.1, the measure $\overline{\lambda}$ that is the restriction of λ^* to the σ-algebra of λ^*-measurable sets is called the **Carathéodory measure induced by λ**.

We now use Theorem 4.3.1 to construct important examples of outer measures.

Example 4.3.3. Let X be a set and $\mathcal{S} = \{\text{finite subset of } X\}$, define $\lambda : \mathcal{S} \to [0,\infty]$ to be $\lambda(A) = |A|^2$. The outer measure λ^* induced by λ counts the element in A. Since any subset A of X satisfies the Carathéodory condition, 2^X is the collection of λ^*-measurable sets, hence $\lambda^* = \overline{\lambda}$ is a measure on 2^X, called counting measure.

This example shows that the Carathéodory measure $\overline{\lambda}$ induced by λ does not necessarily extend λ. □

Example 4.3.4. Let $\mathcal{S} = \{(a,b) : a,b \in [-\infty,\infty], a < b\}$, define $\lambda(a,b) = b - a$ in case $a,b \in \mathbb{R}$, otherwise $\lambda(a,b) = \infty$. Then $\lambda^* = m^*$ is the Lebesgue outer measure, and $\overline{\lambda}$ is the Lebesgue measure. □

Due to outer regularity of Lebesgue measure, consideration of $G_{\sigma\delta}$ sets becomes an indispensable tool in the integration theory on $\mathbb{R}$. We can define an analogue in a general setting:

Definition 4.3.5. Let $\mathcal{S}$ be a collection of subsets of X. We denote $\mathcal{S}_\sigma$ a collection of subsets of X that are union of countably many members in $\mathcal{S}$. We denote $\mathcal{S}_{\sigma\delta}$ the the collection of subsets that are intersection of countably many members of $\mathcal{S}_\sigma$.

Proposition 4.3.6. *Let $\lambda : \mathcal{S} \to [0,\infty]$ be a set function on a collection $\mathcal{S}$ of subsets of X. Let $\overline{\lambda}$ be the Carathéodory measure induced by λ, E a subset of X such that $\lambda^*(E) < \infty$, then there is a subset A of X such that*

$$A \in \mathcal{S}_{\sigma\delta}, \quad E \subseteq A \quad \text{and} \quad \lambda^*(E) = \lambda^*(A).$$

Moreover, if E and each member in $\mathcal{S}$ are λ^*-measurable, then so is A and $\overline{\lambda}(A - E) = 0$.

Proof. We leave the proof as an exercise. The technique has been used many times in measure theory on $\mathbb{R}$. □

4.3.2 Premeaure, Semiring and Extension Theorem

Having the experience on $\mathbb{R}$, we try to imitate the construction of Lebesgue measure in a general setting. To obtain Lebesgue measure on $\mathbb{R}$, we have defined the concepts of length and outer measure. Theorem 4.3.1 guarantees the constructibility of outer measure as long as we have a collection $\mathcal{S}$ and "length" defined on each member of $\mathcal{S}$. And by restricting an outer measure λ^* to the σ-algebra of λ^*-measurable subsets, we get a measure, but the story is not yet complete.

In ideal case the measure $\overline{\lambda}$ should really extend λ. But Example 4.3.3 shows us the Carathéodory measure $\overline{\lambda}$ induced by λ does not necessarily extend λ. This suggests we have to impose finer structures on $\mathcal{S}$ and λ such that $\overline{\lambda}$ actually extends λ. Therefore, there are two problems to be solved:

- When does $\overline{\lambda}$ extend λ?
- If $\overline{\lambda}$ extends λ, when is it unique? i.e., when is $\overline{\lambda}$ the unique measure defined on the σ-algebra of λ^*-measurable sets that extend λ?

Suppose $\lambda : \mathcal{S} \to [0,\infty]$ can be extended to a measure, then by Proposition 4.2.6, λ has to be countably monotone, finitely additive, hence $\lambda(\emptyset) = 0$.

Definition 4.3.7. Let $\mathcal{S}$ be a collection of subsets of X and $\lambda : \mathcal{S} \to [0,\infty]$ a set function, then λ is said to be a **premeasure** if it satisfies the following:

(i) $\lambda(\emptyset) = 0$.

(ii) If $S_i \in \mathcal{S}$ and $\bigsqcup_{i=1}^n S_i \in \mathcal{S}$, then $\lambda(\bigsqcup_{i=1}^n S_i) = \sum_{i=1}^n \lambda(S_i)$.

(iii) If $S, S_1, S_2, \cdots \in \mathcal{S}$ and $S \subseteq \bigcup S_i$, then $\lambda(S) \le \sum \lambda(S_i)$.

Definition 4.3.8. A collection $\mathcal{S}$ of subsets of X is said to be **closed under relative complement** if

$$A, B \in \mathcal{S} \implies A - B \in \mathcal{S}.$$

Note that if $\mathcal{S}$ is closed under relative complement, then $\mathcal{S}$ is closed under intersection since for $A, B \in \mathcal{S}$, $A \cap B = A - (A - B)$.

Now we get a partial solution to question 1.

Theorem 4.3.9. *Let $\mathcal{S}$ be a collection of subsets of X and $\lambda : \mathcal{S} \to [0,\infty]$ a set function. If $\mathcal{S}$ is closed under relative complement, λ is a premeasure, then the Carathéodory measure $\overline{\lambda}$ extends λ.*

Proof. We need to show that each $S \in \mathcal{S}$ is λ^*-measurable and $\overline{\lambda}(S) = \lambda^*(S) = \lambda(S)$. Let's fix $S \in \mathcal{S}$.

For every $A \in X$, we need to verify

$$\lambda^*(A) \ge \lambda^*(A \cap S) + \lambda^*(A - S).$$

It suffices to check those A with $\lambda^*(A) < \infty$. For each $\epsilon > 0$ there are $S_i \in \mathcal{S}$ such that $\bigcup S_i \supseteq A$ and $\lambda^*(A) + \epsilon > \sum \lambda(S_i)$. By assumption $S_i \cap S, S_i - S \in \mathcal{S}$, hence by finite additivity of premeasure, $\lambda(S_i) = \lambda(S_i \cap S) + \lambda(S_i - S)$. Moreover, $\bigcup(S_i \cap S) \supseteq A \cap S$ and $\bigcup(S_i - S) \supseteq A - S$, by definition of outer measure,

$$\lambda^*(A) + \epsilon > \sum \lambda(S_i) = \sum (\lambda(S_i \cap S) + \lambda(S_i - S)) \ge \lambda^*(A \cap S) + \lambda^*(A - S)$$

for all $\epsilon > 0$. We conclude S is λ^*-measurable.

$\lambda(S) = \lambda^*(S)$ is a direct consequence of countable monotonicity. $\square$

The natural collection like intervals on $\mathbb{R}$ is not closed with respect to relative complement, so we need to expand the existing collection $\mathcal{S}$ in order to apply Theorem 4.3.9. Motivated by the collection of intervals on $\mathbb{R}$, one natural choice is to define

$$\mathcal{S}_\sqcup := \left\{ \bigsqcup_{i=1}^{n} S_i : \{S_i \in \mathcal{S}\}_{i=1}^{n} \text{ is a disjoint collection, } n \geq 1 \right\}, \tag{4.3.10}$$

but additional structure of $\mathcal{S}$ must be imposed. To see this, let $\{A_i\}, \{B_i\}$ be finite disjoint collections in $\mathcal{S}$, then

$$\bigsqcup_{i=1}^{m} A_i - \bigsqcup_{i=1}^{n} B_i = \bigsqcup_{i=1}^{m} (A_i - B_1 - \cdots - B_m),$$

we hope successively $A_i - B_1 = \bigsqcup_{\text{finite}} A_i'$, $A_i' - B_2 = \bigsqcup_{\text{finite}} A_i'', \ldots$, where $A_i', A_i'', \cdots \in \mathcal{S}$, so that after finitely many steps, $\bigsqcup_{i=1}^{m} A_i - \bigsqcup_{i=1}^{n} B_i \in \mathcal{S}_\sqcup$, this is the notion of semiring defined below.

Definition 4.3.11. A collection $\mathcal{S}$ of subsets of X is said to be a **semiring** if it satisfies the following:

(i) If $A, B \in \mathcal{S}$, then $A \cap B \in \mathcal{S}$.

(ii) If $A, B \in \mathcal{S}$, then there are $S_1, \ldots, S_n \in \mathcal{S}$ such that $A - B = \bigsqcup_{i=1}^{n} S_i$.

Condition (ii) implies $\emptyset \in \mathcal{S}$ and condition (i) provides us a technical convenience when defining premeasure on $\mathcal{S}_\sqcup$. We shall see this in the next proposition.

Proposition 4.3.12. *Let $\mathcal{S}$ be a semiring of subsets of X. Define $\mathcal{S}_\sqcup$ as in (4.3.10), then*

(i) $\mathcal{S}_\sqcup$ contains $\mathcal{S}$ and is closed under relative complement.

(ii) Any premeasure on $\mathcal{S}$ has a unique extension to a premeasure on $\mathcal{S}_\sqcup$.

Proof. (i) It follows from the discussion preceding Definition 4.3.11.

(ii) We let $\lambda : \mathcal{S} \to [0, \infty]$ be a premeasure. For $E = \bigsqcup_{i=1}^{n} S_i \in \mathcal{S}_\sqcup$, define $\ell(E) = \sum_{i=1}^{n} \lambda(S_i)$. We need to check that ℓ is well-defined. Suppose $\bigsqcup_{i=1}^{n} S_i = E = \bigsqcup_{i=1}^{m} T_i$, for some $T_i \in \mathcal{S}$, since

$$S_i = \bigsqcup_{j=1}^{m} (S_i \cap T_j) \quad \text{and} \quad T_j = \bigsqcup_{i=1}^{n} (S_i \cap T_j),$$

then by finite additivity of premeasure,

$$\sum_{i=1}^{n} \lambda(S_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda(S_i \cap T_j) = \sum_{j=1}^{m} \sum_{i=1}^{n} \lambda(S_i \cap T_j) = \sum_{j=1}^{m} \lambda(T_j).$$

so ℓ is well-defined. We need to show ℓ is finitely additive and countably monotone. The finite additivity of ℓ inherits directly from that of λ. Next, let $A, B_1, B_2, \cdots \in \mathcal{S}_\sqcup$ be such that $A \subseteq \bigcup_{i=1}^{\infty} B_i$. Write $A = \bigsqcup_{i=1}^{n} A_i$, for some $A_i \in \mathcal{S}$ and let

$$\bigcup_j B_j = \underbrace{B_1}_{\bigsqcup_j B_{1j}} \sqcup \bigsqcup_{i=2}^{\infty} \underbrace{(B_i - B_1 - \cdots - B_{i-1})}_{\bigsqcup_j B_{ij}} = \bigsqcup_j B'_j.$$

Then

$$\ell(A) = \sum_{i=1}^{n} \lambda(A_i) \le \sum_{i=1}^{n} \sum_{j=1}^{\infty} \lambda(A_i \cap B'_j) = \sum_{j=1}^{\infty} \ell(A \cap B'_j) \le \sum_{j=1}^{\infty} \ell(B'_j),$$

the last inequality holds since finite additivity implies monotonicity. Since $\sum_{j=1}^{\infty} \ell(B'_j) = \sum_i \left(\sum_j \ell(B_{ij}) \right) \le \sum_i \ell(B_i)$, ℓ is countably monotone.

Finally uniqueness is clear as premeasurs on $\mathcal{S}_\sqcup$ are finitely additive. $\square$

Definition 4.3.13. The set function $\lambda : \mathcal{S} \to [0,\infty]$ is ***σ*-finite** if $X = \bigcup_{k=1}^{\infty} S_k$, where $S_k \in \mathcal{S}$ and $\lambda(S_k) < \infty$.

Now we are in a position to answer questions 1 and 2.

Theorem 4.3.14 (Carathéodory-Hahn). *Let $\lambda : \mathcal{S} \to [0,\infty]$ be a premeasure on a semiring $\mathcal{S}$ of subsets of X.*

(i) The Carathéodory measure $\overline{\lambda}$ induced by λ extends λ.

(ii) If λ is σ-finite, then so is $\overline{\lambda}$ and $\overline{\lambda}$ is the unique measure on the σ-algebra of λ^*-measurable subsets that extends λ.

Proof. (i) By Proposition 4.3.12, the premeasure λ extends to a premeasure λ' on $\mathcal{S}_\sqcup$. By Theorem 4.3.9, λ' induces a Carathéodory measure $\overline{\lambda'}$ that extends λ'. We now show that $(\lambda')^* = \lambda^*$, so that $\overline{\lambda} = \overline{\lambda'}$ extends λ.

Let $E \subseteq X$, we observe that E can be covered by countably many members in $\mathcal{S}$ iff E can be covered by countably many members in $\mathcal{S}_\sqcup$. Let's assume $\bigcup S_i \supseteq E$, for some $S_i \in \mathcal{S}_\sqcup$, we can write $S_i = \bigsqcup_j S_{ij}, S_{ij} \in \mathcal{S}$, so

$$\sum \lambda'(S_i) = \sum \sum_j \lambda(S_{ij}) \ge \lambda^*(E).$$

By taking infimum, $(\lambda')^*(E) \ge \lambda^*(E)$. For the reverse inequality, let $\bigcup S_i \supseteq E$, where $S_i \in \mathcal{S}$, then $\sum \lambda(S_i) = \sum \lambda'(S_i) \ge (\lambda')^*(E)$, so that $\lambda^*(E) \ge (\lambda')^*(E)$.

(ii) $\overline{\lambda}$ is σ-finite as it extends λ. Suppose there is another measure μ defined on the σ-algebra of λ^*-measurable subsets that extends λ. Since there are $X_k \in \mathcal{S}$ such that $X = \bigcup_{k=1}^{\infty} X_k$, $\lambda(X_k) < \infty$. By countable additivity of measures it suffices to check that $\overline{\lambda}$ and μ agree on measurable subsets of X_k, for each k.

Let $A \subseteq X_k$ be λ^*-measurable, then $\overline{\lambda}(A) < \infty$ and thus by Proposition 4.3.6 there is a $\mathcal{S}_{\sigma\delta}$ set $S \supseteq A$ such that $\overline{\lambda}(A) = \overline{\lambda}(S)$. Taking intersection if necessary, we assume $S \subseteq X_k$. Denote $\mathcal{S}^k = \{S \in \mathcal{S} : S \subseteq X_k\}$, by induction $\overline{\lambda}$ and μ agrees on finite union of subsets in $\mathcal{S}^k$. By continuity of measure $\overline{\lambda}$ and μ agrees on $\mathcal{S}^k_\sigma$. Since $\mathcal{S}^k_\sigma$ is closed under finite intersection, by continuity of measure again $\overline{\lambda}$ and μ agree on $\mathcal{S}^k_{\sigma\delta}$. As $S = S \cap X_k$, $S \in \mathcal{S}^k_{\sigma\delta}$, so

$$\overline{\lambda}(A) = \overline{\lambda}(S) = \mu(S) = \mu(A) + \mu(S - A). \tag{4.3.15}$$

As $\overline{\lambda}(S-A) = 0 = \lambda^*(S-A)$, so for each $\epsilon > 0$ there is a $O \in \mathcal{S}_\sigma, O \supseteq S - A$ such that $\epsilon > \mu(O) \geq \mu(S-A)$, thus $\mu(S-A) = 0$. By (4.3.15) we conclude $\overline{\lambda}$ and μ agree on all λ^*-measurable subsets of X_k. $\square$

Let's summarize what we have done so far:

$$\begin{tikzcd}
\text{semiring } \mathcal{S} \text{ and premeasure } \lambda \arrow[r] \arrow[dr, dashed] & \text{induced outer measure } \lambda^* \arrow[d] \\
& \text{Carathéodory measure } \overline{\mu} \text{ induced by } \lambda
\end{tikzcd}$$

The following corollary follows from the technique in the proof of Theorem 4.3.14, we leave it as an exercise.

Corollary 4.3.16. *Let $\mathcal{S}$ be a semiring in X and $\sigma(\mathcal{S})$ the smallest σ-algebra in X containing $\mathcal{S}$. Let μ_1, μ_2 be two measures on $\sigma(\mathcal{S})$ such that $\mu_1|_{\mathcal{S}}$ and $\mu_2|_{\mathcal{S}}$ are σ-finite, then $\mu_1 = \mu_2$ on $\sigma(\mathcal{S})$ iff $\mu_1 = \mu_2$ on $\mathcal{S}$.*

4.3.3 Different Settings for Extension Theorem

It is useful to know the following commonly used terminology.

Definition 4.3.17. Let $\mathcal{S}$ be a collection of subsets of X.

(i) $\mathcal{S}$ is a **ring** if it is closed under finite union and relative complement.

(ii) $\mathcal{S}$ is an **algebra** if it is a ring and $X \in \mathcal{S}$.

(iii) $\mathcal{S}$ is **semialgebra** if it is a semiring and $X \in \mathcal{S}$.

$$\begin{tikzcd}
\sigma\text{-algebra} \arrow[r, hook] & \text{algebra} \arrow[r, hook] \arrow[dr, hook] & \text{ring} \arrow[dr, hook] & \\
& & \text{semialgebra} \arrow[r, hook] & \textbf{semiring}
\end{tikzcd}$$

We can make use of the diagram to show certain collection of subsets is a semiring. The situation is very similar to proving a commutative ring is a UFD, it is sometimes simpler to prove it is an ED (e.g., $F[X]$, F is a field) or a PID (e.g., $F[[X]]$, F is a field) with the help of the additional structure with which an object is endowed.

Up to now we have developed desired extension theorem for several purposes. It is worth noting that there are different approaches and settings to get an extension theorem. Some start with a semialgebra and some start with an algebra[(1)], instead of a semiring. Even the definition of premeasure is different from that defined Definition 4.3.7.

We need to worry about the term "premeasure" when reading other texts. Some books (e.g., [**?**]) replace the conditions (ii) and (iii) in Definition 4.3.7 of premeasure by

$$\bigsqcup_{i=1}^{\infty} A_i \in \mathcal{S} \implies \lambda\left(\bigsqcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \lambda(A_i). \tag{4.3.18}$$

That is, finite additivity and countable monotonicity on $\mathcal{S}$ are replaced by countable additivity on $\mathcal{S}$. It requires little more effort to show these definitions are indeed equivalent when the set function is defined on a semiring. We are going to prove it in Proposition 4.3.22 after Lemma 4.3.19 and Lemma 4.3.20.

Lemma 4.3.19. *Let $\lambda : \mathcal{S} \to [0,\infty]$ be a finitely additive set function on a semiring $\mathcal{S}$. Let $A, A_1, A_2, \ldots, A_n \in \mathcal{S}$ be such that $\bigsqcup_{i=1}^{n} A_i \subseteq A$, then*

$$\sum_{i=1}^{n} \lambda(A_i) \le \lambda(A).$$

Proof. As remarked in the last paragraph right before Definition 4.3.11, there are $S_1, S_2, \ldots, S_m \in \mathcal{S}$ such that $A - A_1 - A_2 - \cdots - A_n = \bigsqcup_{i=1}^{m} S_i$, hence $A = (\bigsqcup_{i=1}^{m} S_i) \sqcup (\bigsqcup_{i=1}^{n} A_i) \in \mathcal{S}$. By finite additivity of λ on $\mathcal{S}$,

$$\lambda(A) = \sum_{i=1}^{m} \lambda(S_i) + \sum_{i=1}^{n} \lambda(A_i) \ge \sum_{i=1}^{n} \lambda(A_i). \qquad \square$$

Lemma 4.3.20. *Let $\lambda : \mathcal{S} \to [0,\infty]$ be a finitely additive set function on a semiring $\mathcal{S}$. Let $A, A_1, A_2, \ldots, A_n \in \mathcal{S}$ be such that $A \subseteq \bigcup_{i=1}^{n} A_i$, then*

$$\lambda(A) \le \sum_{i=1}^{n} \lambda(A_i).$$

Proof. Write $A = \bigcup_{i=1}^{n} (A \cap A_i)$, define

$$A_1' = A \cap A_1 \quad \text{and} \quad A_i' = A \cap A_i - A_1 - \cdots - A_{i-1} \text{ for } i > 1.$$

We can find $S_{i1}, S_{i2}, \ldots, S_{in_i} \in \mathcal{S}$ so that $A_i' = \bigsqcup_{j=1}^{n_i} S_{ij}$, hence $A = \bigsqcup_{i=1}^{n} A_i' = \bigsqcup_{i=1}^{n} \bigsqcup_{j=1}^{n_i} S_{ij}$. Note that $\bigsqcup_{j=1}^{n_i} S_{ij} \subseteq A_i$, hence by Lemma 4.3.19 and finite additivity of λ,

$$\lambda(A) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n_i} \lambda(S_{ij})\right) \le \sum_{i=1}^{n} \lambda(A_i), \tag{4.3.21}$$

as desired. $\square$

[(1)]We have used the idea implicitly in extending the premeasure on a semiring. In the first sentence of the proof of Carathéodory-Hahn Theorem 4.3.14, we extend the premeasure on a semiring $\mathcal{S}$ to a premeasure on $\mathcal{S}_\sqcup$ by Proposition 4.3.12. $\mathcal{S}_\sqcup$ is in fact a ring since it is closed under finite union and relative complement. If we allow $X \in \mathcal{S}_\sqcup$, then $\mathcal{S}_\sqcup$ is an algebra. What is nice in the semiring setting is that semiring needs not contain X so that sometimes we don't need to worry about unbounded length which we often encounter in checking a set function is a premeasure.

Proposition 4.3.22. *Let $\lambda : \mathcal{S} \to [0,\infty]$ be a set function on a semiring $\mathcal{S}$, then λ is finitely additive and countably monotone iff λ is countably additive.*

Proof. Assume λ is finitely additive and countably monotone on $\mathcal{S}$. Since countable monotonicity implies subadditivity, by Lemma 4.3.19 and exactly the same way as Theorem 2.7.1, λ is countably additive.

Conversely, suppose λ is countably additive on $\mathcal{S}$, then it is finitely additive. Let $A, A_1, A_2, \cdots \in \mathcal{S}$ be such that $A \subseteq \bigcup_{i=1}^{\infty} A_i$. Note that we cannot use Lemma 4.3.20, but we can imitate its proof. Since λ is assumed countably additive, (4.3.21) holds even if n is replaced by ∞, so λ is countably monotone. $\square$

4.3.4 Lebesgue-Stieltjes Measure on $\mathbb{R}$

We have seen that Lebesgue measure is a countably additive set function on $\mathbb{R}$, in fact there are other possible choices other than the one induced by length of intervals. As an application of the extension Theorem 4.3.14, we are going to construct them in Proposition 4.3.24.

Note that the collection of left-open, right-closed intervals

$$\mathcal{S} := \{(a,b] : a,b \in \mathbb{R}, a \le b\} \tag{4.3.23}$$

forms a semiring on $\mathbb{R}$. Our aim is to construct a premeasure on it. Now we begin to construct a large family of Borel measures as follows:

Proposition 4.3.24. *Let $F : \mathbb{R} \to \mathbb{R}$ be an increasing function. Let $\mu_F : \mathcal{S} \to [0,\infty)$ be defined by:*

$$\mu_F(a,b] := F(b) - F(a)$$

(i) μ_F is a finitely additive set function on $\mathcal{S}$.

(ii) If F is right-continuous, μ_F is countably additive on $\mathcal{S}$.

Note that by Proposition 4.3.22, (ii) of Proposition 4.3.24 implies μ_F is a premeasure on $\mathcal{S}$, and hence μ_F can be extended to a unique Borel measure on $\mathbb{R}$ because $\mu_F(n,n+1] < \infty$. When $F(x) = x$, μ_F reduces to Lebesgue measure on $\mathbb{R}$.

Proof. (i) Let $(a,b]$ be a finite disjoint union of members in $\mathcal{S}$, we may assume $(a,b] = \bigcup_{i=1}^{n}(a_{i-1},a_i]$ with $a = a_0 < a_1 < a_2 < \cdots < a_n = b$, then

$$\mu_F(a,b] = F(b) - F(a) = \sum_{i=1}^{n}(F(a_i) - F(a_{i-1})) = \sum_{i=1}^{n}\mu_F(a_{i-1},a_i],$$

so μ_F is finitely additive.

(ii) Assume F is right-continuous, let $\{(a_i,b_i]\}$ be a countable disjoint collection of intervals in $\mathcal{S}$ such that $(a,b] = \bigsqcup_{i=1}^{\infty}(a_i,b_i] \in \mathcal{S}$. By Lemma 4.3.19 since $\bigsqcup_{i=1}^{n}(a_i,b_i] \subseteq (a,b]$ for each n, we have

$$\sum_{i=1}^{n}\mu_F(a_i,b_i] \le \mu_F(a,b] \text{ for each } n \implies \sum_{i=1}^{\infty}\mu_F(a_i,b_i] \le \mu_F(a,b].$$

To prove the reverse inequality, let $\epsilon > 0$ be given, then by right-continuity we can find a $\delta > 0$ such that

$$\mu_F(a,b] - \mu_F(a+\delta,b] < \epsilon. \tag{4.3.25}$$

Now $(a+\delta,b] \subseteq [a+\delta,b] \subseteq \bigsqcup_{i=1}^{\infty}(a_i,b_i]$, again by right-continuity we can find $\delta_i > 0$ such that $F(b_i+\delta_i) - F(b_i) < \epsilon/2^i$. Now

$$[a+\delta,b] \subseteq \bigsqcup_{i=1}^{\infty}(a_i,b_i+\delta_i) \implies [a+\delta,b] \subseteq \bigsqcup_{i=1}^{n}(a_i,b_i+\delta_i)$$

for some n, and hence $(a+\delta,b] \subseteq \bigsqcup_{i=1}^{n}(a_i,b_i+\delta_i]$. So by Lemma 4.3.20,

$$\begin{aligned}\mu_F(a+\delta,b] &\le \sum_{i=1}^{n}\mu_F(a_i,b_i+\delta_i] \le \sum_{i=1}^{\infty}(F(b_i+\delta_i)-F(a_i)) \\ &< \sum_{i=1}^{\infty}(F(b_i)-F(a_i)+\epsilon/2^i) = \sum_{i=1}^{\infty}(F(b_i)-F(a_i))+\epsilon.\end{aligned}$$

Combining with (4.3.25), $\mu_F(a,b] < \sum_{i=1}^{\infty}\mu_F(a_i,b_i]+2\epsilon$, for each $\epsilon > 0$. □

We have shown that given an increasing function, we get a finitely additive set function μ_F. Further, if it is right-continuous, then we get a measure μ_F on $\mathbb{R}$. In fact the converse is also true!

Proposition 4.3.26.

(i) Let $\mu : \mathcal{S} \to [0,\infty)$ be a finitely additive set function such that $\mu(a,b] < \infty$ for every $a,b \in \mathbb{R}$, then there exists an increasing function $F : \mathbb{R} \to \mathbb{R}$ such that $\mu(a,b] = F(b) - f(a)$.

(ii) If μ is also countably additive, then F is right-continuous.

Proof. (i) Suppose such F exists, then by fixing $a = 0$ and letting b vary, we have $\mu(0,b] = F(b) - F(0)$, which motivates the following definition:

$$F(x) := \begin{cases} \mu(0,x], & x > 0, \\ 0, & x = 0, \\ -\mu(x,0], & x < 0, \end{cases}$$

then $\mu(a,b] = F(b) - F(a)$.

(ii) Let $x_n > x$ decreases to x, then since $(x,x_1] = \bigsqcup_{k=2}^{\infty}(x_k,x_{k-1}]$, by countable additivity we have

$$F(x_1) - F(x) = \sum_{k=2}^{\infty}(F(x_{k-1}) - F(x_k)) = \lim_{n\to\infty}(F(x_1) - F(x_n)),$$

hence $F(x) = \lim_{n\to\infty} F(x_n)$. □

Let $\mathcal{S}$ be defined as in (4.3.23), F be increasing and right-continuous and let μ_F^* denote the outer measure induced by $\mu_F : \mathcal{S} \to [0,\infty)$. Let Σ_F denote the collection of μ_F^*-measurable subsets of $\mathbb{R}$. Since the measure on Σ_F that extends μ_F is always unique, we shall always denote $\mu_F = \mu_F^*|_{\Sigma_F}$ and bear in mind that μ_F is a complete measure defined on $\Sigma_F \supset \mathcal{B}_{\mathbb{R}}$. For each $A \in \Sigma_F$,

$$\mu_F(A) = \inf\left\{\sum_{i=1}^{\infty}\mu_F(a_i,b_i] : \bigcup_{i=1}^{\infty}(a_i,b_i] \supseteq A\right\}.$$

We now show that those left-open, right-closed intervals can be replaced by open intervals, and then we show that μ_F also enjoys some useful regular properties as in Lebesgue measure.

Lemma 4.3.27. *For each $A \in \Sigma_F$,*

$$\mu_F(A) = \inf\left\{\sum_{i=1}^{\infty}\mu_F(a_i,b_i) : \bigcup_{i=1}^{\infty}(a_i,b_i) \supseteq A\right\}. \tag{4.3.28}$$

Proof. Denote the RHS of (4.3.28) by $\mu(A)$. Let $\{(a_i,b_i)\}_{i=1}^{\infty}$ be an open cover of A, then

$$\mu_F(A) \le \sum_{i=1}^{\infty}\mu_F(a_i,b_i),$$

hence $\mu_F(A) \le \mu(A)$. Conversely, let $\bigcup_{i=1}^{\infty}(a_i,b_i] \supseteq A$. For each i we may find $\delta_i > 0$ such that $\mu_F(a_i,b_i+\delta_i) - \mu_F(a_i,b_i] < \epsilon/2^i$, then

$$\mu(A) \le \sum_{i=1}^{\infty}\mu_F(a_i,b_i+\delta_i) < \sum_{i=1}^{\infty}\mu_F(a_i,b_i] + \epsilon$$

for each $\epsilon > 0$, so $\mu(A) \le \sum_{i=1}^{\infty}\mu_F(a_i,b_i]$ and hence $\mu(A) \le \mu_F(A)$. □

Theorem 4.3.29. *For each $A \in \Sigma_F$,*

$$\mu_F(A) = \inf\{\mu_F(U) : U \supseteq A,\ U \text{ open}\} \tag{4.3.30}$$
$$= \sup\{\mu_F(K) : K \subseteq A,\ K \text{ compact}\}. \tag{4.3.31}$$

Proof. Let U be open, and $U \supseteq A$, then as μ_F is a measure, $\mu_F(U) \ge \mu_F(A)$. By Lemma 4.3.27 for each $\epsilon > 0$ we can find an open cover $\{(a_i,b_i)\}$ of A such that $\sum\mu_F(a_i,b_i) \le \mu_F(A)+\epsilon$. Let $U = \bigcup(a_i,b_i)$, then U is open and $\mu_F(U) \le \mu_F(A)+\epsilon$, so (4.3.30) holds.

Assume first that A is bounded, then so is $\overline{A}$. To do inner approximation of A, we do outer approximation of its complement relative to a larger set. Let $\epsilon > 0$ be given, by (4.3.30) we can find an open set U such that $U \supseteq \overline{A}-A$ and $\mu_F(U) < \mu_F(\overline{A}-A)+\epsilon$, then $\overline{A}-U \subseteq A$ is compact and

$$\mu_F(A)-\mu_F(\overline{A}-U) = \mu_F(A\cap U) = \mu_F(U)-\mu_F(U-A) \le \mu_F(U)-\mu_F(\overline{A}-A) < \epsilon.$$

Together with $\mu_F(K) \le \mu_F(A)$ for each compact subset K of A we have shown that (4.3.31) holds when A is bounded. For general $A \in \Sigma_F$ let $\epsilon > 0$ be given and let $A_n = A\cap(n-1,n]$. Then for each n we can find compact $K_n \subseteq A_n$ such that $\mu_F(A_n) - \mu_F(K_n) < \epsilon/2^{|n|}$. Now for each N,

$$\mu_F\left(\bigcup_{n=-N}^{N} A_n\right) \le 3\epsilon + \mu_F\left(\bigcup_{n=-N}^{N} K_n\right),$$

hence by taking $N \to \infty$, we are done. □

The last two theorems are analogous version of theorems on Lebesgue measure. We leave the proofs as exercises for readers.

Theorem 4.3.32. *Let $A \subseteq \mathbb{R}$, then the following are equivalent:*

(i) $A \in \Sigma_F$.

(ii) $A = G - N_1$, *where* G *is* G_δ *and* $\mu_F(N_1) = 0$.

(iii) $A = F \cup N_2$, *where* F *is* F_σ *and* $\mu(N_2) = 0$.

Theorem 4.3.33. *If* $A \in \Sigma_F$ *and* $\mu(E) < \infty$, *then for every* $\epsilon > 0$ *there are open intervals* $I_1, I_2, \ldots, I_n$ *such that* $\mu_F(A \Delta \bigsqcup_{i=1}^n I_i) < \epsilon$.

4.4 Exercises and Problems

Exercises

4.1. Show that the set functions defined in Example 4.2.4 and Example 4.2.5 are measures.

4.2. Suppose $X \to X$ is an invertible map such that $S \in \mathcal{S}$ iff $\phi(S) \in \mathcal{S}$ and $\lambda(\phi(S)) = \lambda(S)$. Prove that the outer measure μ^* induced by λ satisfies $\mu^*(\phi(A)) = \mu^*(A)$.

4.3. In Example 4.2.15 we have seen that $(\mathbb{R}, \mathcal{B}, \mu)$ is not complete. Show that $(\mathbb{R}, \mathcal{L}, m)$ is its completion.

4.4. Prove Proposition 4.2.19.

4.5. Prove Corollary 4.3.16.

4.6. Prove that given a collection $\mathcal{S}$ of subsets X and (i) of Definition 4.3.11 holds, then (ii) in Definition 4.3.11 of semiring:

> "If $A, B \in \mathcal{S}$, then there are $S_1, \ldots, S_n \in \mathcal{S}$ such that $A - B = \bigsqcup_{i=1}^n S_i$."

is equivalent to

> "If $A, A_1 \in \mathcal{S}$ such that $A_1 \subseteq A$, then there is a disjoint collection $\{A_k \subseteq X - A_1\}_{k=2}^n$ in $\mathcal{S}$ such that $A = \bigsqcup_{k=1}^n A_k$."

4.7. Let $(X, \mathcal{T}_X)$ be a topological space. For a subset Y of X, let $\mathcal{T}_Y$ denote the subspace topology of Y induced by X and let $\sigma(\mathcal{T}_X), \sigma(\mathcal{T}_Y)$ denote the Borel σ-algebra on X and Y respectively. Show that

$$\sigma(\mathcal{T}_Y) = \sigma(\mathcal{T}_X) \cap Y := \{U \cap Y : U \in \sigma(\mathcal{T}_X)\}.$$

4.8. Let $f : X \to Y$ be a function and Σ a σ-algebra on Y, show that

$$f^{-1}(\Sigma) := \{f^{-1}(A) : A \in \Sigma\}$$

is also a σ-algebra.

4.9. If $f : X \to Y$ is a function between two sets and $\mathcal{S}$ is a nonempty collection of subsets of Y, then

$$\sigma(f^{-1}(\mathcal{S})) = f^{-1}(\sigma(\mathcal{S})).$$

4.10. Let X be a topological space, show that

$$\mathcal{S} := \{L \cap U : L \text{ is closed, } U \text{ is open}\} = \{A - B : A, B \text{ are closed}\}$$

forms a semiring of subsets of X.

4.11. Prove that an intersection of semirings is not necessarily a semiring. Give an example by considering $X := \{1,2,3\}$.

4.12. Let $\mathcal{S} = \{(a,b),(a,b],[a,b),[a,b] : a,b \in \mathbb{R}, a \leq b\}$. By definition, $\mathcal{S}$ contains $\{\emptyset\}$ and $\{\{a\} : a \in \mathbb{R}\}$. Show that each of the following collections is a semiring:

(i) $\mathcal{S}$ itself.

(ii) $\mathcal{S} \times \mathcal{S} \subseteq 2^{\mathbb{R}^2}$ defined by $\{S_1 \times S_2 : S_1, S_2 \in \mathcal{S}\}$.

(iii) The n-fold products of $\mathcal{S}$, i.e., $\mathcal{S}^n := \underbrace{\mathcal{S} \times \cdots \times \mathcal{S}}_{n \text{ times}} \subseteq 2^{\mathbb{R}^n}$.

4.13. **(Generalize Problem 6.3.28)** Let $\mathcal{S}$ and $\mathcal{T}$ be semirings of subsets of X and Y respectively. Then $\mathcal{S} \times \mathcal{T} := \{S \times T : S \in \mathcal{S}, T \in \mathcal{T}\}$ is a semiring of subsets of $X \times Y$. We call $\mathcal{S} \times \mathcal{T}$ a **product semiring**.

4.14. Show that a collection $\mathcal{S}$ of subsets of X is a semialgebra iff the following holds:

(i) $\emptyset, X \in \mathcal{S}$.

(ii) If $A, B \in \mathcal{S}$, $A \cap B \in \mathcal{S}$

(iii) If $A \in \mathcal{S}$, $X - A$ is a finite disjoint union of members in $\mathcal{S}$.

4.15. Let $X = \mathbb{Q}$, $\mathcal{S} = \{(a,b] \cap \mathbb{Q} : a \leq b\}$ and $\mathcal{S}_\cup = \{\bigcup_{i=1}^n S_i : S_i \in \mathcal{S}, n \geq 1\}$. Define $\lambda(a,b] = \infty$ if $a < b$ and $\lambda(\emptyset) = 0$ if $a = b$.

(a) Show that $\mathcal{S}$ is closed under relative complement and $\lambda : \mathcal{S} \to [0,\infty]$ is a premasure.

(b) Show that the extension of λ to the smallest σ-algebra containing $\mathcal{S}_\cup$ is not unique.

This problem tells us σ-finiteness in (ii) of Carathéodory-Hahn theorem cannot be dropped.

4.16. Let F be increasing and right-continuous and let μ_F be the Lebesgue-Stieltjes measure induced by F. Show that

$$\begin{aligned}
\mu_F(\{a\}) &= F(a) - F(a^-),\\
\mu_F[a,b) &= F(b^-) - F(a^-),\\
\mu_F[a,b] &= F(b) - F(a^-),\\
\mu_F(a,b) &= F(b^-) - F(a).
\end{aligned}$$

4.17. Prove Theorem 4.3.32.

4.18. Prove Theorem 4.3.33.

Chapter 5

Measurable Functions and Integration

We have mentioned in Chapter 3 that measurable functions are the natural class of functions for which we can do another type of integration analogous to the Riemann one. Namely, we want to approximate measurable functions by simple functions and define the integral of measurable functions in terms of simple functions.

In this chapter Σ denotes a σ-algebra on a space X and μ denotes a measure on Σ. We will not mention it in each of the results. Some results do not require a measure, as indicated in the statement.

5.1 Measurable Functions

Many results concerning Lebesgue measurable functions can be translated directly to measurable ones with respect to a σ-algebra on a space X. However, since it is too restrictive to assume the measure space to be complete, changes have to be made.

In the past for $E \in \mathcal{L}$, on $(E, \mathcal{L} \cap E, m)$ we say that a property $P(x)$ holds a.e. on E if $m\{x \in E : P(x) \text{ doesn't hold}\} = 0$ (Definition 3.1.7). This definition works very well for complete measure space, but not for incomplete ones because we don't even know whether or not $\{P(x) \text{ doesn't hold}\}$ is measurable! We need to reformulate our notion of "almost everywhere" so that the concept of "negligible sets" can be carried to the study of general measure space.

Definition 5.1.1. Let (X, Σ, μ) be a measure space. If there is a set $X_0 \in \Sigma$ such that a property related to points $x \in X$ holds on $X - X_0$ and $\mu(X_0) = 0$, then we say that the property holds **almost everywhere (abbr. a.e.) on X** or that the property holds **for almost every (abbr. a.e.) x on X**.

The proofs from Proposition 5.1.2 to simple approximation Theorem 5.1.12 are all almost identical to those in Chapter 3, we leave them as exercises and don't repeat the proofs all over again.

Proposition 5.1.2. *Let f be an extended real-valued function defined on X. Then the following statements are equivalent:*

(i) For each $c \in \mathbb{R}$, $\{x \in E : f(x) > c\}$ is measurable.

(ii) *For each $c \in \mathbb{R}$, $\{x \in E : f(x) \geq c\}$ is measurable.*

(iii) *For each $c \in \mathbb{R}$, $\{x \in E : f(x) < c\}$ is measurable.*

(iv) *For each $c \in \mathbb{R}$, $\{x \in E : f(x) \leq c\}$ is measurable.*

Any one of the above statements implies for each extended real number c, $f^{-1}(c)$ is measurable.

Definition 5.1.3. Let (X,Σ) be a measurable space. A function f is **measurable** if it is extended real-valued and one of the 4 conditions in Proposition 5.1.2 holds.

As mentioned in Chapter 3 measurable functions is a class of functions whom we can approximate using simple functions. Measurability of f is a key point to construct such simple functions (recall the proof of Lemma 3.3.4).

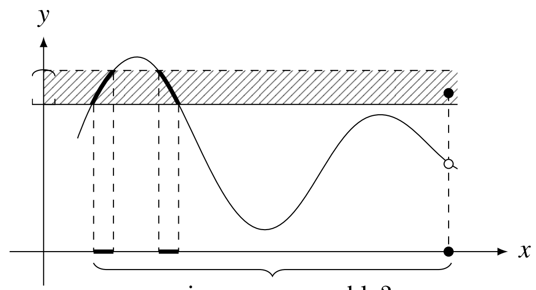


Figure 5.1: Preimage of a measurable function.

Until we arrive to Definition 5.2.37, by measurable functions we mean extended real-valued measurable functions. To emphasize a measurable function f is real-valued, we say f is real-valued measurable function.

Proposition 5.1.4. *Let f be a real-valued function on X. Then f is measurable iff for each open $O \subseteq \mathbb{R}$, $f^{-1}(O)$ is measurable.*

Proposition 5.1.5. *Let $\{f_k\}_{k=1}^n$ be a finite family of measurable functions on a common domain $E \in \Sigma$, the functions $\max_{1\leq k\leq n} f_k$ and $\min_{1\leq k\leq n} f_k$ are measurable.*

Hence as remarked before, the functions

$$f^+ := \max\{f(x),0\}, \quad f^- := \max\{-f(x),0\}$$

are both nonnegative measurable and $f = f^+ - f^-$.

Proposition 5.1.6. *Let (X,Σ,μ) be complete, $X_0 \in \Sigma$ a set with $\mu(X - X_0) = 0$, then an extended real-valued function f on X is measurable iff $f|_{X_0}$ is measurable.*

As a consequence, if X is complete and f,g are extended real valued functions such that $f = g$ a.e., then f is measurable on X iff g is measurable on X.

Remark. Proposition 5.1.6 can be false if X is incomplete. For example, if there is $E \notin \Sigma$ but $E \subseteq Z \in \Sigma$, $\mu(Z) = 0$, then $0 = \chi_E$ except possibly on Z.

Proposition 5.1.7. *Let f,g be measurable real-valued functions on X. Then:*

(i) For each $\alpha,\beta \in \mathbb{R}$, $\alpha f + \beta g$ is measurable.

(ii) $f \cdot g$ is measurable.

Proposition 5.1.8. *Let f be a measurable real-valued function on X and $g : \mathbb{R} \to \mathbb{R}$ continuous, then the composition $g \circ f : X \to \mathbb{R}$ is measurable.*

Proposition 5.1.9. *Let $\{f_n\}$ be a sequence of measurable functions on X such that $f_n \to f$ pointwise a.e. on X. If (X,Σ,μ) is complete, or the convergence is pointwise on all of X, then f is measurable.*

Remark. Proposition 5.1.9 can be false if X is incomplete, can you give an example? We have discussed some examples of incomplete measure spaces.

Proposition 5.1.10. *Let $\{f_n\}$ be a sequence of measurable functions on X, then the following functions are measurable*

$$\sup_{n\geq 1} f_n, \quad \inf_{n\geq 1} f_n, \quad \overline{\lim_{n\to\infty}} f_n \quad \text{and} \quad \underline{\lim_{n\to\infty}} f_n.$$

Lemma 5.1.11 (Simple Approximation). *Let f be a measurable real-valued function on $E \in \Sigma$. Assume f is bounded on E, then for each $\epsilon > 0$, there are simple functions φ_ϵ and ψ_ϵ defined on E such that*

$$\varphi_\epsilon \leq f \leq \psi_\epsilon \quad \text{and} \quad 0 \leq \psi_\epsilon - \varphi_\epsilon < \epsilon \text{ on } E.$$

Remark. The simple approximation lemma tells us every bounded measurable function is a uniform limit of a sequence of simple functions. Also from the proof of the lemma when f is nonnegative, we can choose $\varphi_\epsilon \geq 0$.

Theorem 5.1.12 (Simple Approximation). *A function f on X is measurable if and only if there is a sequence of simple functions $\{\varphi_n\}$ on X for which $\varphi_n \to f$ pointwise on X and*

$$|\varphi_n| \leq |f| \text{ on } X \text{ for all } n.$$

(i) If X is σ-finite, we can further assume each φ_n vanishes outside a set of finite measure. We describe such functions have finite support.

(ii) If f is nonnegative, we may choose $\{\varphi_n\}$ to be increasing with $\varphi_n \geq 0$.

Theorem 5.1.13 (Egoroff). *Suppose $\mu(X) < \infty$ and $f, f_1, f_2, \ldots$ are measurable real-valued functions on X such that $f_n \to f$ a.e., then for every $\epsilon > 0$, there is $E \subseteq X$ such that $f_n \rightrightarrows f$ on E and $\mu(X - E) < \epsilon$.*

Loosely put, $f_n \rightrightarrows f$ on $E \in \Sigma$ iff we can find a sequence of positive integers $\{n_k\}$ such that for each $x \in E$, for any k and for each $m \geq n_k$, $|f_m(x) - f(x)| < \frac{1}{k}$, that said, iff

$$E \subseteq \bigcap_{k=1}^{\infty} \bigcap_{m=n_k}^{\infty} \left\{ x \in X : |f_m(x) - f(x)| < \frac{1}{k} \right\}.$$

If RHS can be constructed such that its complement has μ-measure less than ϵ, we may take E to be RHS. To construct RHS, it is same as constructing its complement

$$\bigcup_{k=1}^{\infty}\bigcup_{m=n_k}^{\infty}\left\{x \in X : |f_m(x)-f(x)| \geq \frac{1}{k}\right\}.$$

Proof. WLOG let's assume $f_n \to f$ pointwise on X. Define $E_n(k) = \bigcup_{m=n}^{\infty}\{x \in X : |f_m(x)-f(x)| \geq \frac{1}{k}\}$. By definition, $E_n(k)$ is descending in n, $\bigcap_{n=1}^{\infty} E_n(k) = \emptyset$ and $\mu(X) < \infty$, by continuity of measure $\lim_{n\to\infty}\mu(E_n(k)) = 0$. So for given $\epsilon > 0$, we can choose n_k large enough so that $\mu(E_{n_k}(k)) < \epsilon/2^k$. Define $X - E = \bigcup_{k=1}^{\infty} E_{n_k}(k)$, we have $\mu(X-E) < \epsilon$, and $f_n \rightrightarrows f$ on E by construction. □

Remark. Egoroff's theorem is also true when $f, f_1, f_2, \ldots$ are complex measurable functions defined in Definition 5.2.37.

5.2 Integration

In summer 2010-2011 we didn't have enough time to do integration theory. It is a chance to present all important results on Lebesgue integration in general setting.

In Chapter 3 we notice that each measurable function f can be written as a difference of nonnegative measurable functions, i.e., $f = f^+ - f^-$. To define integration of a measurable function, it suffices to do so for nonnegative measurable ones.

5.2.1 Integration of Nonnegative Functions

Definition 5.2.1. Let ϕ be a nonnegative simple function, say $\phi = \sum_{i=1}^{n} a_i \chi_{A_i}$, $a_i \in \mathbb{R}, A_i$'s $\in \Sigma$ are disjoint and partition X, we define

$$\int_X \phi\, d\mu = \sum_{i=1}^{n} a_i \mu(A_i).$$

Note that $\int_X \phi\, d\mu$ takes value in $[0,\infty]$. We need to check that the integral in Definition 5.2.1 is well-defined. To do this, assume $\sum_{i=1}^{m} a_i \chi_{A_i} = \phi = \sum_{j=1}^{n} b_j \chi_{B_j}$, $\bigsqcup_{i=1}^{m} A_i = X = \bigsqcup_{j=1}^{n} B_j$, then

$$\begin{aligned}\sum_{i=1}^{m} a_i \mu(A_i) &= \sum_{i=1}^{m} a_i \sum_{j=1}^{n} \mu(A_i \cap B_j) = \sum_{j=1}^{n}\sum_{i=1}^{m} a_i \mu(A_i \cap B_j)\\ &= \sum_{j=1}^{n} \sum_{i, A_i\cap B_j \neq \emptyset} b_j \mu(A_i \cap B_j) = \sum_{j=1}^{n} b_j \sum_{i=1}^{m} \mu(A_i \cap B_j) = \sum_{j=1}^{n} b_j \mu(B_j),\end{aligned}$$

showing that $\int_E \phi\, d\mu$ is well-defined, so is the following definition.

Definition 5.2.2. If ϕ is a nonnegative simple function on X and $E \in \Sigma$, define

$$\int_E \phi\, d\mu = \int_X \chi_E \phi\, d\mu.$$

Having the definition of integration of simple functions, one may, traditionally, define the integration of measurable function $f : X \to [0,\infty]$ over X by

$$\int_X f\,d\mu = \sup_{\substack{0\le\phi\le f\\ \phi \text{ simple}}} \int_X \phi\,d\mu. \tag{5.2.3}$$

However, this definition seems weird at the beginning because we are just doing "inner approximation". In order to convince ourself this is a suitable definition, we define integration of measurable functions in another way under which some basic properties of integral can be easily verified. Our goal is to show that our choice of integration of f in Definition 5.2.6 takes the same value as (5.2.3). Before showing our definition is well-defined, we need some preliminary results.

Proposition 5.2.4. *Let ϕ,ϕ_1 and ϕ_2 be nonnegative simple functions and $\alpha \ge 0$, then:*

(i) $0 \le \int_X \phi\,d\mu \le \infty$.

(ii) $\int_X \alpha\phi\,d\mu = \alpha\int_X \phi\,d\mu$.

(iii) If $\phi_1 \ge \phi_2$ *on* X, $\int_X \phi_1\,d\mu \ge \int_X \phi_2\,d\mu$.

(iv) $\int_X(\phi_1+\phi_2)\,d\mu = \int_X \phi_1\,d\mu + \int_X \phi_2\,d\mu$.

(v) $\nu(E) := \int_E \phi\,d\mu$ *is a measure on* Σ. *If* $\mu(E)=0$, *then* $\nu(E)=0$.

Proof. (i), (ii) They are immediately true by definition.

(iii), (iv) We write $\phi_1 = \sum_{i=1}^m a_i\chi_{A_i}$ and $\phi_2 = \sum_{j=1}^n b_j\chi_{B_j}$, then (iii), (iv) follows from the representations

$$\phi_1-\phi_2 = \sum_{i=1}^m\sum_{j=1}^n (a_i-b_j)\chi(A_i\cap B_j) \quad\text{and}\quad \phi_1+\phi_2 = \sum_{i=1}^m\sum_{j=1}^n (a_i+b_j)\chi(A_i\cap B_j).$$

(v) Let $\phi = \sum_{i=1}^n c_i\phi_{C_i}$ and $E = \bigsqcup_{k=1}^\infty E_k, E_k \in \Sigma$. The only nontrivial property to check is countable additivity, but

$$\nu(E) = \int_X \sum_{i=1}^n c_i\chi_{E\cap C_i}\,d\mu = \sum_{i=1}^n c_i\mu(E\cap C_i) = \sum_{k=1}^\infty \int_{E_k}\sum_{i=1}^n c_i\chi_{C_i}\,d\mu = \sum_{k=1}^\infty \nu(E_k),$$

so ν is a measure. It is clear that $\mu(E)=0$ implies $\nu(E)=0$. □

Proposition 5.2.5. *Let $\{\phi_n\}$ be an increasing sequence of nonnegative simple functions such that $\phi_n \to \phi$ pointwise on X, for some nonnegative simple function ϕ, then*

$$\int_X \phi\,d\mu = \lim_{n\to\infty}\int_X \phi_n\,d\mu$$

This is a special case of monotone convergence theorem. The technique in this proof will be used again to prove the general case, as long as the integration of a measurable function can be defined.

Proof. Since $\phi_n \le \phi$, for all n, one has $\lim \int_X \phi_n \, d\mu \le \int_X \phi \, d\mu$. To show the reverse inequality, fix $c \in (0,1)$, then $X_n := \{x \in X : c\phi \le \phi_n\} \in \Sigma$ is an ascending collection with $\bigcup_{n=1}^{\infty} X_n = X$, hence

$$c \int_X \phi \, d\mu = \lim \int_{X_n} c\phi \, d\mu \le \lim \int_{X_n} \phi_n \, d\mu \le \lim \int_X \phi_n \, d\mu,$$

the first equality used the fact that $A \mapsto \int_A \phi \, d\mu$ is a measure. As $c \in (0,1)$ is arbitrary, we obtain $\int_X \phi \, d\mu \le \lim \int_X \phi_n \, d\mu$. □

Now we can define our integration:

Definition 5.2.6 (Version 1). Let $f : X \to [0,\infty]$ be measurable, then there is an increasing sequence of nonnegative simple functions $\{\phi_n\}$ such that $\phi_n \to f$ pointwise on X, we define the (Lebesgue) integral over X by

$$\int_X f \, d\mu = \lim_{n\to\infty} \int_X \phi_n \, d\mu.$$

Different from the definition of integral in (5.2.3), the measurability of f is obviously important for the integral defined in Definition 5.2.6.

Proposition 5.2.7. *The integration in Definition 5.2.6 is well-defined.*

We introduce some useful notations. For $a,b \in \mathbb{R}$, we define $a \vee b = \max\{a,b\}$ and $a \wedge b = \min\{a,b\}$[(1)]. For functions $f,g : X \to [-\infty,\infty]$, we define a function $f \vee g$ pointwise by $(f \vee g)(x) = f(x) \vee g(x)$. $f \wedge g$ is defined similarly.

Proof. Let $\{\phi_n\}$ and $\{\phi'_n\}$ be two increasing sequences of nonnegative simple functions, $\phi_n, \phi'_n \to f$ pointwise on X. We need to show $\lim_{n\to\infty} \int_X \phi_n \, d\mu = \lim_{n\to\infty} \int_X \phi'_n \, d\mu$. Fix an $m \in \mathbb{N}$, clearly $\{\phi'_m \wedge \phi_n\}_{n=1}^{\infty}$ is increasing and converges to ϕ'_m pointwise on X, hence by Proposition 5.2.5,

$$\int_X \phi'_m \, d\mu = \lim_{n\to\infty} \int_X \phi'_m \wedge \phi_n \, d\mu \le \lim_{n\to\infty} \int_X \phi_n \, d\mu.$$

But m is arbitrary, so $\lim_{n\to\infty} \int_X \phi'_n \, d\mu \le \lim_{n\to\infty} \int_X \phi_n \, d\mu$. Reversing the roles of ϕ_n and ϕ'_n, we get the reverse inequality, so that Definition 5.2.6 is well-defined. □

Now we can justify (5.2.3) is a suitable definition of integration.

Proposition 5.2.8. *Let $f : X \to [0,\infty]$ be measurable, then*

$$\int_X f \, d\mu = \sup_{\substack{0 \le \phi \le f \\ \phi \text{ simple}}} \int_X \phi \, d\mu. \tag{5.2.9}$$

Proof. Denote RHS of (5.2.9) by β. Let $\{\phi_n\}$ be an increasing sequence of nonnegative simple functions, $\phi_n \nearrow f$. By definition $\int_X f \, d\mu = \lim \int_X \phi_n \, d\mu \le \beta$. Next there are two ways to show $\beta \le \int_X f \, d\mu$.

[(1)]To memorize them, recall that the operation $\cup$ always enlarges a set and $\cap$ always shrinks a set.

Method 1. β being a supremum, there are nonnegative simple $\psi_n \le f$ such that $\beta = \lim_{n\to\infty} \int_X \psi_n \, d\mu$. Fix an n and integrate both sides of $\psi_n \le \psi_n \vee \phi_m$, one has

$$\int_X \psi_n \, d\mu \le \int_X \psi_n \vee \phi_m \, d\mu \le \lim_{m\to\infty} \int_X \psi_n \vee \phi_m \, d\mu = \int_X f \, d\mu.$$

We relax the choice of n so that $\beta \le \int_X f \, d\mu$.

Method 2. Fix a nonnegative simple $\psi \le f$ and fix a $c \in (0,1)$. Let $X_n = \{x \in X : c\psi \le \phi_n\}$, it is easy to see $\{X_n\}$ is ascending and $X = \bigcup X_n$, hence

$$c \int_X \psi \, d\mu = \int_{\bigcup X_n} c\psi \, d\mu = \lim \int_{X_n} c\psi \, d\mu \le \lim \int_{X_n} \phi_n \, d\mu \le \lim \int_X \phi_n \, d\mu = \int_X f \, d\mu,$$

i.e., $c \int_X \psi \, d\mu \le \int_X f \, d\mu$. As the choice of c can be relaxed, $\int_X \psi \, d\mu \le \int_X f \, d\mu$, also the choice of ψ can be relaxed, $\beta \le \int_X f \, d\mu$.

3rd direct proof. Define ψ_n as in method 1, then we construct an increasing sequence of simple functions by $\varphi_n = \psi_1 \vee \psi_2 \vee \cdots \vee \psi_n \vee \phi_n$[(2)]. Then

$$\beta \leftarrow \int_X \psi_n \, d\mu \le \int_X \varphi_n \, d\mu \le \beta, \quad f(x) \leftarrow \phi_n(x) \le \varphi_n(x) \le f(x)$$

implies $\beta = \lim \int_X \varphi_n \, d\mu = \int_X f \, d\mu$. □

Henceforth we have two (equivalent) definitions of integration, they are used interchangeably to deal with different situations.

Definition 5.2.10 (Version 2). Let $f : X \to [0,\infty]$ be measurable, the (Lebesgue) integral of f over X is

$$\int_X f \, d\mu = \sup_{\substack{0\le\phi\le f \\ \phi \text{ simple}}} \int_E \phi \, d\mu,$$

and for each $E \in \Sigma$, we define the integration of f over E by

$$\int_E f \, d\mu = \int_X \chi_E f \, d\mu.$$

Now we extend some basic properties of integration of nonnegative measurable functions listed in Proposition 5.2.4, except for (v) which we will prove very soon.

Proposition 5.2.11. *Let f,g be nonnegative measurable functions and $\alpha \ge 0$, then:*

(i) $0 \le \int_X f \, d\mu \le \infty$.

(ii) $\int_X \alpha f \, d\mu = \alpha \int_X f \, d\mu$.

(iii) If $f \ge g$ *on* X, $\int_X f \, d\mu \ge \int_X g \, d\mu$.

(iv) $\int_X (f+g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu$

(v) $\nu(A) := \int_A f \, d\mu$ *is a finitely additive set function on* Σ.

---

(2) If $a_1 \le b_1, a_2 \le b_2$, then $a_1 \vee b_1 \le a_2 \vee b_2$, take $a_1 = \psi_1 \vee \cdots \vee \psi_n$, $a_2 = \psi_1 \vee \cdots \vee \psi_{n+1}$, $b_1 = \phi_n$, $b_2 = \phi_{n+1}$, so that $\{\varphi_n\}$ is increasing.

Proof. (i) It is immediately true.

(iii) Let's fix a nonnegative simple function ϕ with $\phi \le g$. Then since $\phi \le f$, $\int_X \phi\, d\mu \le \int_X f\, d\mu$. Since the choice of ϕ can be relaxed, $\int_X g\, d\mu \le \int_X f\, d\mu$.

(ii), (iv) Let's choose increasing sequences of nonnegative simple functions $\{\phi_n\}$ and $\{\varphi_n\}$ such that $\phi_n \to f$ and $\varphi_n \to g$ pointwise on X. For (ii),

$$\int_X \alpha f\, d\mu = \lim \int_X \alpha \phi_n\, d\mu = \alpha \lim \int_X \phi_n\, d\mu = \alpha \int_X f\, d\mu.$$

For (iv), by (iv) of Proposition 5.2.4,

$$\int_X (f+g)\, d\mu = \lim \int_X (\phi_n + \varphi_n)\, d\mu = \lim \left(\int_X \phi_n\, d\mu + \int_X \varphi_n\, d\mu \right).$$

(v) Let $A = \bigsqcup_{k=1}^n A_k$, then $\nu(A) = \int_X \chi_A f\, d\mu = \int_X \sum_{k=1}^n \chi_{A_k} f\, d\mu$, and the result follows from (iv). □

Next we can discuss the interesting part of the integration theory. They tell us limit operations are easier to handle in Lebesgue integration.

Theorem 5.2.12 (Lebesgue's Monotone Convergence). *Let $\{f_n\}$ be a sequence of nonnegative measurable functions on X. If $\{f_n\}$ is increasing and $f_n \to f$ pointwise on X, then f is measurable and*

$$\lim_{n\to\infty} \int_X f_n\, d\mu = \int_X f\, d\mu.$$

We just need to modify the proof of Proposition 5.2.5. Note that a crucial step is to apply (v) of Proposition 5.2.4, we try to "shrink" a simple $\phi \le f$ by a multiplicative constant.

Proof. The measurability of f is guaranteed by Proposition 5.1.9. By (iii) of Proposition 5.2.11, $\lim \int_X f_n\, d\mu \le \int_X f\, d\mu$.

To prove the reverse inequality, let's fix a nonnegative simple function $\phi \le f$ and $c \in (0,1)$. Consider $X_n := \{x \in X : c\phi \le f_n\} \in \Sigma$, it is ascending and $\bigcup_{n=1}^\infty X_n = X$. Thus by (v) of Proposition 5.2.4,

$$\int_X c\phi\, d\mu = \lim \int_{X_n} c\phi\, d\mu \le \lim \int_{X_n} f_n\, d\mu \le \lim \int_X f_n\, d\mu.$$

By relaxing the choice of c, $\int_X \phi\, d\mu \le \lim \int_X f_n\, d\mu$. Finally, by relaxing the choice of ϕ, $\int_X f\, d\mu \le \lim \int_X f_n\, d\mu$. □

Theorem 5.2.13. *Let $f_n : X \to [0,\infty]$ be measurable for $n = 1,2,\ldots$, then*

$$\int_X \sum_{n=1}^\infty f_n\, d\mu = \sum_{n=1}^\infty \int_X f_n\, d\mu.$$

Proof. Let $g_N = \sum_{n=1}^N f_n$, then by (iv) of Proposition 5.2.11, $\int_X g_N\, d\mu = \sum_{n=1}^N \int_X f_n(x)\, d\mu$. By monotone convergence theorem,

$$\int_X \sum_{n=1}^\infty f_n(x)\, d\mu = \int_X \lim g_N\, d\mu = \lim \int_X g_N\, d\mu = \sum_{n=1}^\infty \int_X f_n(x)\, d\mu. \qquad \square$$

Lemma 5.2.14 (Fatou). *If $\{f_n\}$ is a sequence of nonnegative measurable functions, then*

$$\int_X \varliminf_{n\to\infty} f_n \, d\mu \le \varliminf_{n\to\infty} \int_X f_n \, d\mu.$$

Proof. Define $g_n = \inf_{i\ge n} f_i$, then $\lim g_n = \varliminf f_n$. As $\{g_n\}$ is an increasing sequence of nonnegative measurable functions, by monotone convergence theorem,

$$\int_X \varliminf f_n \, d\mu = \int_X \lim g_n \, d\mu = \lim \int_X g_n \, d\mu \le \varliminf \int_X f_n \, d\mu. \qquad \square$$

Remark. The inequality cannot be reversed in general. To see this, let $X = [0,1]$, $f_n = nx^{n-1}$ and $\mu = m$ be Lebesgue measure. Then $\lim_{n\to\infty} f_n = 0$ a.e. and $\int_{[0,1]} f_n \, dm = 1$ (by Theorem 5.2.21), hence $\int_{[0,1]} \varliminf f_n \, dm = 0 < 1 = \varliminf \int_{[0,1]} f_n \, dm$.

Now we extend property (v) of Proposition 5.2.4.

Theorem 5.2.15. *Suppose $f : X \to [0,\infty]$ is measurable, then*

$$\nu(E) := \int_E f \, d\mu$$

is a measure on Σ. Moreover, for every measurable $g : X \to [0,\infty]$, we have

$$\int_X g \, d\nu = \int_X g f \, d\mu.$$

Proof. To show ν is a measure, it suffices to show ν is countably additive. Let $E = \bigsqcup_{k=1}^{\infty} E_k$, then $\chi_E = \sum_{k=1}^{\infty} \chi_{E_k}$, thus

$$\nu(E) = \int_E f \, d\mu = \int_X \chi_E f \, d\mu = \int_X \sum_{k=1}^{\infty} \chi_{E_k} f \, d\mu = \sum_{k=1}^{\infty} \nu(E_k),$$

so ν is indeed countably additive.

Next by simple approximation theorem, there is an increasing sequence $\{\phi_n\}$ of nonnegative simple functions such that $\phi_n \to g$ pointwise on X. Combined with monotone convergence theorem and linearity of integration, it suffices to check the last assertion for $g = \chi_A$, for $A \in \Sigma$, which is obvious. $\square$

5.2.2 Integration of General Measurable Functions

Let $f : X \to [0,\infty]$ be measurable, then f^+, f^- are nonnegative measurable functions for which $\int_X f^+ \, d\mu$ and $\int_X f^- \, d\mu$ make sense. Owing to the equality $f = f^+ - f^-$ it is reasonable to define $\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu$. However, problem arises in the case $\int_X f^+ \, d\mu = \int_X f^- \, d\mu = +\infty$, we introduce the following definition to overcome this difficulty:

Definition 5.2.16. A function $f : X \to [-\infty,\infty]$ is said to be **integrable over *X* with respect to $\boldsymbol{\mu}$** (or simply **integrable**) if it is measurable and both $\int_X f^+ \, d\mu, \int_X f^- \, d\mu$ are finite, moreover, the integral of f over X is defined by

$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu.$$

Since $|f| = f^+ + f^-$, one can easily verify that f is integrable iff f^+ and f^- are integrable iff $|f|$ is integrable.

There are some usual notations for collection of integrable functions.

Definition 5.2.17. For $1 \le p < \infty$, we denote by $\mathcal{L}^p(X,\mu)$ (or by $\mathcal{L}^p(X)$ or $\mathcal{L}^p(\mu)$) the collection of all integrable functions f such that $\int_X |f|^p \, d\mu < \infty$. When $p = \infty$, we denote by $\mathcal{L}^\infty(X,\mu)$ the collection of **essentially bounded** functions[(3)]. For $f,g \in \mathcal{L}^p(X,\mu)$, we define an equivalence relation $\sim$ by $f \sim g$ iff $f = g$ almost everywhere. Then we define

$$L^p(X) = \mathcal{L}^p(X)/\sim.$$

It is customary to write $f \in L^p(X)$ instead of $[f] \in L^p(X)$ and for $f,g \in L^p(X)$, the notation $f = g$ is understood as $f = g$ almost everywhere.

We now extend the list of properties listed in Proposition 5.2.11.

Proposition 5.2.18. *Let f and g be integrable over X, $\alpha \in \mathbb{R}$, then:*

(i) $\int_X \alpha f \, d\mu = \alpha \int_X f \, d\mu$.

(ii) $\int_X (f+g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu$

(iii) If $f \ge g$ on X, $\int_X f \, d\mu \ge \int_X g \, d\mu$.

(iv) $\left| \int_X f \, d\mu \right| \le \int_X |f| \, d\mu$.

(v) $\nu(A) := \int_A f \, d\mu$ is a countably additive set function on Σ.

Proof. (i) Observe that

$$(\alpha f)^+ = \begin{cases} \alpha f^+, & \text{if } \alpha \ge 0, \\ -\alpha f^-, & \text{if } \alpha < 0, \end{cases} \quad \text{and} \quad (\alpha f)^- = \begin{cases} \alpha f^-, & \text{if } \alpha \ge 0, \\ -\alpha f^+, & \text{if } \alpha < 0. \end{cases}$$

(i) follows from considering the cases that $\alpha \ge 0$ and $\alpha < 0$.

(ii) We have proved (ii) is true when f,g are nonnegative. Now we prove the general case. Since f,g are integrable, they are finite a.e. (why?), so we let X_0 be such that f,g are finite on $X - X_0$ with $\mu(X_0) = 0$. Then $(f+g)^+ - (f+g)^- = f + g = f^+ - f^- + g^+ - g^-$ on $X - X_0$ implies

$$(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+$$

on $X - X_0$. Since each term is nonnegative, we integrate both sides over $X - X_0$ and rearrange terms, so that (ii) is proved.

(iii) Since $f - g \ge 0$, $\int_X (f-g) \, d\mu \ge 0$ by definition, hence the result follows from (ii).

(iv) By definition of integral of general measurable functions,

$$\left| \int_X f \, d\mu \right| = \left| \int_X f^+ \, d\mu - \int_X f^- \, d\mu \right| \le \int_X f^+ \, d\mu + \int_X f^- \, d\mu = \int_X |f| \, d\mu.$$

(v) Since $\int_A f \, d\mu = \int_A f^+ \, d\mu - \int_A f^- \, d\mu$, each term is countably additive by Theorem 5.2.15 and everything converges absolutely, so ν is countably additive. $\square$

[(3)] i.e., collection of measurable functions f such that there is $C > 0$, $\mu\{x \in X : |f| > C\} = 0$.

Next, as we notice, whenever a set function is countably additive on Σ we have continuity of measure on Σ for ascending collection $\{X_n \in \Sigma\}$, we also have that for descending collection $\{Y_n \in \Sigma\}$ if there is Y_k that has finite measure. Since by measure we mean a nonnegative and countably additive set function, we modify the name of our observation a little bit.

Theorem 5.2.19 (Continuity of Integration). *Let f be integrable over X.*

(i) If $\{X_n \in \Sigma\}$ is ascending, then

$$\int_{\bigcup_{n=1}^{\infty} X_n} f \, d\mu = \lim_{n\to\infty} \int_{X_n} f \, d\mu.$$

(ii) If $\{X_n \in \Sigma\}$ is descending, then

$$\int_{\bigcap_{n=1}^{\infty} X_n} f \, d\mu = \lim_{n\to\infty} \int_{X_n} f \, d\mu.$$

Proof. The proof is identical to Theorem 2.7.3. □

Theorem 5.2.20 (Lebesgue's Dominated Convergence). *Let $f, f_1, f_2, \ldots$ be measurable functions on X such that $f_n \to f$ pointwise a.e. on X. If there is a function g integrable over X such that for each $n \in \mathbb{N}$,*

$$|f_n| \le g \text{ a.e. on } X,$$

then f is integrable over X and

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Proof. Since

$$\int_X |f| \, d\mu = \int_X \lim |f_n| \, d\mu \le \underline{\lim} \int_X |f_n| \, d\mu \le \int_X g \, d\mu,$$

f is integrable over X. Moreover, since $g - f_n, g + f_n \ge 0$, by Fatou's lemma,

$$\int_X \underline{\lim}(g - f_n) \, d\mu \le \underline{\lim} \int_X (g - f_n) \, d\mu \implies \overline{\lim} \int_X f_n \, d\mu \le \int_X f \, d\mu,$$

$$\int_X \underline{\lim}(g + f_n) \, d\mu \le \underline{\lim} \int_X (g + f_n) \, d\mu \implies \int_X f \, d\mu \le \underline{\lim} \int_X f_n \, d\mu,$$

so $\lim \int_X f_n$ exists and is equal to $\int_X f \, d\mu$. □

5.2.3 Relation Between the Lebesgue and Riemann Integrals

In this subsection we would like to explain why Lebesgue integration theory "extends" the notion of Riemann integrals (strictly speaking, Lebesgue integral extends the Riemann integral of the class of absolutely Riemann integrable functions). After that we will see how superior the Lebesgue theory is in handling limit operations.

Throughout this subsection we let $\mu = m$, as before, denote Lebesgue measure on $\mathbb{R}$. The Lebesgue integral of f over $[a,b]$ is denoted by $\int_{[a,b]} f \, dm$, and the Riemann one is denoted by $\int_a^b f \, dx$.

Theorem 5.2.21. *Let $f : [a,b] \to \mathbb{R}$ be a Riemann integrable function, then $f \in \mathcal{L}^1([a,b],m)$, moreover,*

$$\int_{[a,b]} f\,dm = \int_a^b f\,dx.$$

Proof. By the meaning of $f \in \mathcal{L}^1([a,b],m)$, we need to show f is measurable and Lebesgue integrable, what's more, two integrals are equal.

Let $\{P_n\}$ be a collection of partitions of $[a,b]$, say $P_n = \{x_{n,0},\dots,x_{n,k_n} : a = x_{n,0} < x_{n,1} < \cdots < x_{n,k_n} = b\}$, we define $\|P_n\| = \max_{1\le i\le k_n} |x_{n,i} - x_{n,i-1}|$. Also we construct the following step functions

$$\varphi_n(x) = \sum_{i=1}^{k_n} \Big(\inf_{t\in[x_{n,i-1},x_{n,i}]} f(t) \Big) \chi_{[x_{n,i-1},x_{n,i})}(x)$$

and

$$\psi_n(x) = \sum_{i=1}^{k_n} \Big(\sup_{t\in[x_{n,i-1},x_{n,i}]} f(t) \Big) \chi_{[x_{n,i-1},x_{n,i})}(x),$$

then by Riemann integrability, whenever $\|P_n\| \to 0$, one has

$$\lim \int_a^b \varphi_n\,dx = \int_a^b f\,dx = \lim \int_a^b \psi_n\,dx. \tag{5.2.22}$$

Now we choose a sequence of partition such that P_{n+1} refines P_n and $\|P_n\| \to 0$, one such possible choice is to divide each subintervals into half. Then the limit in (5.2.22) is achieved, moreover, $\{\varphi_n\}$ is increasing[(4)] and $\{\psi_n\}$ is decreasing with

$$\varphi_n(x) \le f(x) \le \psi_n(x)$$

for each $x \in [a,b)$. Let $\varphi(x) = \lim_{n\to\infty} \varphi_n(x)$ and $\psi = \lim_{n\to\infty} \psi_n(x)$, both limits exist since f is bounded. Then

$$\int_{[a,b]} (\psi - \varphi)\,dm \le \int_{[a,b]} (\psi_n - \varphi_n)\,dm = \int_a^b \psi_n\,dx - \int_a^b \varphi_n\,dx$$

for each n, and hence by (5.2.22), $\int_{[a,b]}(\psi - \varphi)\,dm = 0$. But $\psi - \varphi \ge 0$, hence $\varphi = \psi$ a.e. on $[a,b]$. As $\varphi \le f \le \psi$, we conclude $f = \lim_{n\to\infty} \varphi_n$ pointwise a.e. on $[a,b]$, and hence measurable. We also let $X_0 \subseteq \mathbb{R}$ be such that $\varphi_n \to f$ pointwise on $[a,b] - X_0$ and $m(X_0) = 0$.

f is Lebesgue integrable because Riemann integrable functions are bounded. By dominated convergence theorem and (5.2.22),

$$\begin{aligned}\int_{[a,b]} f\,dm &= \int_{[a,b]-X_0} f\,dm = \lim \int_{[a,b]-X_0} \varphi_n\,dm \\ &= \lim \int_{[a,b]} \varphi_n\,dm = \lim \int_a^b \varphi_n\,dx = \int_a^b f\,dx.\end{aligned}$$

□

(4) For example, let $x \in [a,b)$, then for each n there is a unique interval $I_n = [x_{n,M_n}, x_{n,M_n+1})$ such that $x \in I_n$. Then $\varphi_n(x) = \inf f(\overline{I_n})$. As P_{n+1} refines P_n, $I_{n+1} \subseteq I_n$, hence $\varphi_{n+1}(x) = \inf \varphi(\overline{I_{n+1}}) \ge \inf \varphi(\overline{I_n}) = \varphi_n(x)$.

Remark. In Theorem 5.2.21 Riemann integrability over $[a,b]$ cannot be changed to improper Riemann integrability over $[a,b]$. To see this, consider

$$g(x) := \sum_{n=1}^{\infty} (-1)^{n-1} n \chi_{(1/(n+1),1/n]}(x),$$

then g is improper Riemann integrable but not Lebesgue integrable.

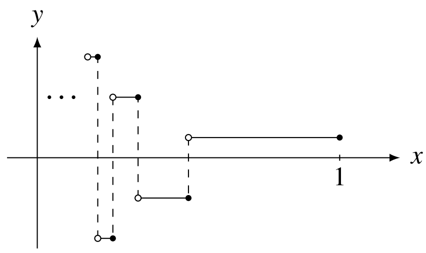


Figure 5.2: Riemann but not Lebesgue integrable.

In the proof of Theorem 5.2.21, we have shown that $\lim_{n\to\infty} \varphi_n = \lim_{n\to\infty} \psi_n$ on $[a,b]$ except $X_0 \subseteq [a,b]$ with $m(X_0) = 0$. Which actually shows that:

Corollary 5.2.23. *Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable, then almost every $x \in [a,b]$ is a point of continuity of f.*

Proof. We adopt all notations in the proof of Theorem 5.2.21. Let $P = \bigcup_{n=1}^{\infty} P_n$ be the collection of all partition points, which is countable. Fix an $x_0 \in X - X_0 - P$, by construction $\lim_{n\to\infty} \varphi_n(x_0) = f(x_0) = \lim_{n\to\infty} \psi_n(x_0)$. So for each $\epsilon > 0$, there is an N so that

$$\psi_N(x_0) - \varphi_N(x_0) < \epsilon.$$

As $x_0 \in X - P$, i.e., $x_0 \notin P_N$, so x_0 must be an interior point of some interval $I := (x_{N,i}, x_{N,i+1})$.

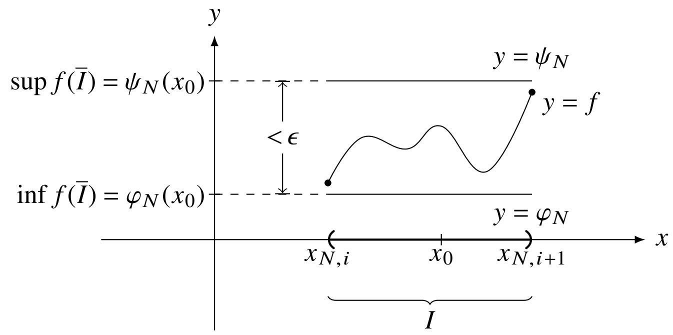


Figure 5.3: How f is bounded by φ and ψ.

For all $x \in I$, both $f(x), f(x_0) \in [\inf f(I), \sup f(I)] \subseteq [\varphi_N(x_0), \psi_N(x_0)]$, and hence

$$|f(x) - f(x_0)| \leq \psi_N(x_0) - \varphi_N(x_0) < \epsilon.$$

That is, x_0 is a point of continuity. So f is continuous on $[a,b]$ except possibly on $X_0 \cup P$ which has Lebesgue measure zero. We conclude f is continuous a.e.. $\square$

As you might have seen somewhere, the converse of Corollary 5.2.23 is also true if f is bounded, whose proof is divided into several steps in Problem 5.11.

Example 5.2.24. [(5)] We try to show that

$$\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{1}{n(n+1)\cdots(n+k)} = \int_0^1 \frac{e^x-1}{x}\,dx.$$

The series converges since for $k = 1,2$, $\frac{1}{n(n+1)}, \frac{1}{n(n+1)(n+2)} < \frac{1}{n^2}$ and for $k \geq 3$, $\frac{1}{n(n+1)\cdots(n+k)} < \frac{1}{n^2k^2}$. Now

$$\frac{1}{x(x+1)\cdots(x+k)} = \sum_{r=0}^{k}\frac{a_r}{x+r} \implies 1 = \sum_{r=0}^{k} a_r \prod_{\substack{0\leq j\leq k\\ j\neq r}}(x+j),$$

by choosing suitable x we can deduce that $1 = a_r(-1)^r r!(k-r)!$, i.e., $a_r = \frac{(-1)^r}{k!}\binom{k}{r}$, which implies

$$\frac{1}{n(n+1)\cdots(n+k)} = \frac{1}{k!}\sum_{r=0}^{k}\binom{k}{r}\frac{(-1)^r}{n+r}. \tag{5.2.25}$$

Since desired answer is an integral, to this end by binomial expansion and integration,

$$\sum_{r=0}^{k}\binom{k}{r}\frac{(-1)^r}{n+r} = (-1)^n\int_0^{-1}(1+x)^k x^{n-1}\,dx = \int_0^1 (1+x)^k x^{n-1}\,dx = \int_0^1 x^k(1-x)^{n-1}\,dx. \tag{5.2.26}$$

Combining (5.2.25) and (5.2.26), we have

$$\begin{aligned}
\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{1}{n(n+1)\cdots(n+k)} &= \sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{1}{k!}\left(\sum_{r=0}^{k}\binom{k}{r}\frac{(-1)^r}{n+r}\right)\\
&= \sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\int_0^1\frac{1}{k!}x^k(1-x)^{n-1}\,dx\\
&= \sum_{n=1}^{\infty}\int_0^1\sum_{k=1}^{\infty}\frac{1}{k!}x^k(1-x)^{n-1}\,dx \qquad (5.2.27)\\
&= \sum_{n=1}^{\infty}\int_0^1(e^x-1)(1-x)^{n-1}\,dx\\
&= \int_0^1\sum_{n=1}^{\infty}(e^x-1)(1-x)^{n-1}\,dx. \qquad (5.2.28)
\end{aligned}$$

(5.2.27) and (5.2.28) are true by monotone convergence theorem, where we have changed dx to dm and then back to dx implicitly. $\square$

[(5)] It is a problem found in mathlinks: `http://www.artofproblemsolving.com/Forum/viewtopic.php?f=67&t=275547`.

It is important to study the Riemann integral of a function over an unbounded interval. Having learnt Theorem 5.2.21, we are also interested in when the improper Riemann integral $\int_a^\infty f(x)\,dx$ of a function $f:[a,\infty)\to\mathbb{R}$ agree with $\int_{[a,\infty)} f\,dm$.

Since Lebesgue integrability (including integration with respect to general measures) is a kind of absolute convergence, some of improper Riemann integrable functions turn out to be not Lebesgue integrable, as indicated in the following well-known example.

Example 5.2.29. Consider

$$f(x) := \sum_{n=1}^{\infty} (-1)^{n-1}\frac{1}{n}\chi_{[n-1,n)}(x).$$

Since $\int_{[0,\infty)} |f|\,dm = \sum_{n=1}^\infty \frac{1}{n} = \infty$, f is not Lebesgue integrable.

Next we show that f is improper Riemann integrable. It is Riemann integrable over all interval $[0,a]$, $a>0$. Moreover,

$$\int_0^\infty f\,dx = \sum_{n=1}^\infty (-1)^{n-1}\frac{1}{n} = \ln 2.$$

To obtain this sum, note that for each $x\in[0,1)$, $\sum_{n=1}^\infty (-1)^{n-1}x^{n-1} = \frac{1}{x+1}$ and the convergence is absolutely, hence the value of the sum is independent of any rearrangement. By the following pairing,

$$g_N(x) = \sum_{n=1}^{2N} (-1)^{n-1}x^{n-1} = (1-x)+(x^2-x^3)+\cdots+(x^{2N-2}-x^{2N-1}),$$

$\{g_N\}$ is increasing, and $\lim g_N = \frac{1}{x+1}$, hence by monotone convergence theorem,

$$\sum_{n=1}^\infty (-1)^{n-1}\frac{1}{n} = \lim\int_{[0,1)} g_N\,dm = \int_{[0,1)} \lim g_N\,dm = \int_0^1 \frac{1}{1+x}\,dx = \ln 2. \qquad \square$$

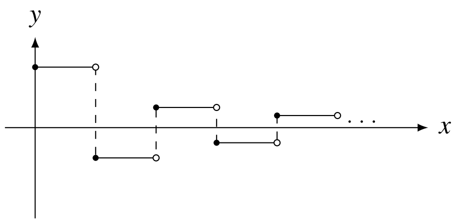


Figure 5.4: f is Riemann integrable due to cancellation.

However, if the improper Riemann integrability is also absolute, then two types of integrals agree.

Theorem 5.2.30. *Let $f:[a,\infty)\to\mathbb{R}$ be Riemann integrable over $[a,b]$, for each $b>a$*[6]*. Then f is Lebesgue integrable iff the improper Riemann integral $\int_a^\infty |f(x)|\,dx$ exists, and in that case,*

$$\int_{[a,\infty)} |f|\,dm = \int_a^\infty |f|\,dx$$

[6] It is also called locally integrable.

and

$$\int_{[a,\infty)} f\,dm = \int_a^\infty f\,dx. \tag{5.2.31}$$

Proof. f is Lebesgue measurable because for each $c \in \mathbb{R}$ (pick $N > a$),

$$f^{-1}(c,\infty) = \bigcup_{n=N}^{\infty} \left((f^{-1}(c,\infty)) \cap [a,n]\right) = \bigcup_{n=N}^{\infty} (f|_{[a,n]})^{-1}(c,\infty),$$

so $f^{-1}(c,\infty)$ is measurable by Riemann integrability over $[a,n]$ for each n.

Assume f is Lebesgue integrable, then $\int_{[a,\infty)} |f|\,dm$ exists. Let $a_1, a_2, \cdots \in [a,\infty)$ be such that $a_n \to \infty$, then by dominated convergence theorem and Theorem 5.2.21,

$$\int_{[a,\infty)} |f|\,dm = \int_{[a,\infty)} \lim \chi_{[a,a_n]}|f|\,dm = \lim \int_{[a,\infty)} \chi_{[a,a_n]}|f|\,dm = \lim \int_a^{a_n} |f|\,dx, \tag{5.2.32}$$

as the choice of a_n's are arbitrary, hence $\int_a^\infty |f|\,dx$ exists.

Conversely, if $\int_a^\infty |f|\,dx$ exists, then we pick a strictly increasing sequence $a_1, a_2 \cdots \in [a,\infty)$ such that $a_n \to \infty$, then each equality in (5.2.32) (from right to left) is true, where the second equality follows form monotone convergence theorem. Hence f is Lebesgue integrable.

Finally we consider the last two equalities. The first equality follows also from (5.2.32). For the second one, let $a_n > a$ with $a_n \to \infty$. Then define $f_n(x) = \chi_{[a,a_n]}(x) f(x)$, by dominated convergence theorem,

$$\int_{[a,\infty)} f\,dm = \int_{[a,\infty)} \lim f_n\,dm = \lim \int_{[a,\infty)} f_n\,dm = \lim \int_a^{a_n} f\,dx.$$

As the choice of a_n's are arbitrary, hence the improper Riemann integral exists, (5.2.31) follows. □

Recall that a function $f : [a,\infty) \to \mathbb{R}$ is said to be absolutely integrable in Riemann sense if f is locally Riemann integrable (then so is $|f|$) and $\int_a^\infty |f|\,dx$ exists.

Example 5.2.33 (Riemann-Lebesgue lemma). Let f be absolutely Riemann integrable on $\mathbb{R}$, we can show that

$$\lim_{\lambda\to\infty} \int_{-\infty}^{\infty} f(t)\cos(\lambda t)\,dt = \lim_{\lambda\to\infty} \int_{-\infty}^{\infty} f(t)\sin(\lambda t)\,dt = 0.$$

We outline the proof briefly without technical detail. It is clear f is measurable on $\mathbb{R}$. The Riemann integral can be switched to Lebesgue integral. We just discuss the integral with cos, the one with sin is essentially the same. Let $\epsilon > 0$ be given, there is an $n \in \mathbb{N}$ such that

$$\left| \left(\int_{\mathbb{R}} - \int_{[-n,n]} \right) f(t)\cos(\lambda t)\,dt \right| < \epsilon. \tag{5.2.34}$$

By simple approximation theorem and Theorem 3.4.1, one can find a step function ϕ on $[-n,n]$ such that $\int_{[-n,n]} |f - \phi|\,dm < \epsilon$, combing this inequality with (5.2.34) one has

$$\left| \int_{\mathbb{R}} f\cos(\lambda t)\,dt - \int_{-n}^{n} \phi\cos(\lambda t)\,dt \right| < 2\epsilon.$$

Since ϕ is a step function, as $\lambda \to \infty$, $\int_{-n}^{n} \phi \cos(\lambda t)\, dt \to 0$. That is to say, whenever λ is large enough,

$$\left| \int_{\mathbb{R}} f \cos(\lambda t)\, dt \right| < 3\epsilon,$$

as desired. You can fill the gaps by Problem 5.28. □

When studying improper or proper Riemann integrals one might have tried the following operations

$$\frac{d}{dy} \int_{?} f(x,y)\, dx = \int_{?} \frac{\partial}{\partial y} f(x,y)\, dx \tag{5.2.35}$$

in order to compute certain kind of integrals. This is a very useful trick since some of the undesired factor in the integrand can be eliminated! For instance, a usual example or exercise in complex analysis is to evalute the following integral

$$\int_0^\infty \frac{\sin x}{x}\, dx.$$

If we introduce the (convergence) factor e^{-ax} $(a > 0)$ to the integrand, then after differentiating $\int_0^\infty e^{-ax} \frac{\sin x}{x}\, dx$ with respect to a, we can get rid of x in the denominator and $\int_0^\infty e^{-ax} \sin x$ is easy to compute (integration by parts twice or evaluate the complex integral $\int e^{-ax} \sin x\, dx = \operatorname{Im} \int e^{-ax} e^{ix}\, dx$).

We will see that dominated convergence theorem is enough to justify such kind of operations in (5.2.35). To state the result precisely, we use the notation defined in Definition 5.3.7. For $(x,t) \in X \times \mathbb{R}$, we define $f_x(t) = f(x,t)$ and $f^t(x) = f(x,t)$.

Theorem 5.2.36. *Let $f : X \times (a,b) \to \mathbb{R}$ satisfy the following conditions:*

(i) For each $t \in (a,b)$, $f^t \in \mathcal{L}^1(X,\mu)$.

(ii) For each $t_0 \in (a,b)$, there is a $\delta > 0$ and a $g \in L^1(X,\mu)$ such that for a.e. x, $f_x(t)$ is differentiable and $|\frac{\partial f}{\partial t}(x,t)| \le g(x)$ on $(t_0 - \delta, t_0 + \delta)$.

Then $\int_X f(x,t)\, d\mu(x)$ is differentiable, moreover,

$$\frac{d}{dt} \int_X f(x,t)\, d\mu(x) = \int_X \frac{\partial f}{\partial t}(x,t)\, d\mu(x).$$

Proof. Let $t \neq t_0 \in (a,b)$, write

$$D(t) = \frac{\int_X f(x,t)\, d\mu(x) - \int_X f(x,t_0)\, d\mu(x)}{t - t_0} = \int_X \frac{f(x,t) - f(x,t_0)}{t - t_0}\, d\mu(x).$$

By definition, there is $\delta > 0$ such that for a.e. x, f_x is differentiable and $|\frac{df_x}{dt}(t)| \le g(x)$ on $(t_0 - \delta, t_0 + \delta)$. Let $t = t_n$ be a sequence such that $t_n \to t_0$. Define $\varphi_n(x) = \frac{f(x,t_n) - f(x,t_0)}{t_n - t_0}$, then $D(t_n) = \int_X \varphi_n(x)\, d\mu$. By mean-value theorem, for large enough n and a.e. x, $|\varphi_n(x)| \le g(x)$. Since $\varphi_n(x) \to \frac{\partial f}{\partial t}(x,t_0)$ pointwise a.e., hence by dominated convergence theorem,

$$\lim D(t_n) = \lim \int_X \varphi_n(x)\, d\mu = \int_X \lim \varphi_n(x)\, d\mu = \int_X \frac{\partial f}{\partial t}(x,t_0)\, d\mu.$$

Since the limit $\lim D(t_n)$ exists for every sequence $\{t_n\}$ such that $t_n \to t_0$, the limit $\lim_{t\to t_0} D(t)$ exists, and we conclude for each $t \in (a,b)$,

$$\frac{d}{dt}\left(\int_X f(x,t)\,d\mu(x)\right)\bigg|_{t=t_0} = \int_X \frac{\partial f}{\partial t}(x,t_0)\,d\mu(x). \qquad \square$$

It is left as an exercise to use Theorem 5.2.36 to evaluate $\int_0^\infty \frac{\sin x}{x}\,dx$.

5.2.4 Integration of Complex Functions

For completeness we also introduce complex-valued measurable functions. They arise very naturally. For example, in the study of pointwise convergence of Fourier series, $S(f)$, of a Riemann integrable function $f : [-\pi,\pi] \to \mathbb{R}$, integration of complex functions provides a succinct way to express a Fourier series:

$$\begin{aligned} S(f) &:= \frac{a_0}{2} + \sum_{n=1}^{\infty}\big(a_n \cos nx + b_n \sin nx\big) \\ &= \sum_{n\in\mathbb{Z}} \underbrace{\left(\frac{1}{2\pi}\int_{[-\pi,\pi)} f(\theta)e^{-in\theta}\,dm(\theta)\right)}_{:=\hat{f}(n)} e^{inx} = \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx}, \end{aligned}$$

where

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\cos n\theta\,d\theta, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\sin n\theta\,d\theta.$$

To study the pointwise convergence, we study the partial sum

$$S_N(f) := \sum_{|n|\le N} \hat{f}(n)e^{inx} = \frac{1}{2\pi}\int_{[-\pi,\pi)} f(\theta) \underbrace{\sum_{n=-N}^{N} e^{in(x-\theta)}}_{=\frac{\sin((N+\frac{1}{2})(x-\theta))}{\sin(\frac{1}{2}(x-\theta))}} dm(\theta)$$

and the difference $|f - S_N(f)|$. The theory can be developed with the ease of handling limit in Lebesgue integration, more specifically, by the complex version of Lebesgue dominated convergence theorem.

Now recall that at the beginning of this chapter, we have fixed the symbols X,Σ and μ to be the components of the measure space (X,Σ,μ), the convention carries over to this subsection.

Definition 5.2.37. A complex-valued function $f : X \to \mathbb{C}$ is said to be **measurable** if $\operatorname{Re} f, \operatorname{Im} f : X \to \mathbb{R}$ are measurable functions.

Now we have extended our class of measurable functions. We need to be careful in using the terms "measurable". If we want to emphasize a measurable function is extended real-valued, we use the term **extended real-valued measurable function** for clarity, likewise we would use the terms **real-valued measurable function** and **complex measurable function**.

The following proposition gives another characterization of the measurability of $f : X \to \mathbb{C}$, which is an analogue of Proposition 5.1.4.

Proposition 5.2.38. *A function $f : X \to \mathbb{C}$ is measurable iff $f^{-1}(B) \in \Sigma$ for each Borel set B in $\mathbb{C}$.*

The proof is left as an exercise. Note that if $\mathbb{C}$ is replaced by $\mathbb{R}$ in Proposition 5.2.38, the same σ-algebra technique also implies $f : X \to \mathbb{R}$ is measurable iff $f^{-1}(B) \in \Sigma$ for each Borel set B in $\mathbb{R}$. A general definition of measurability of function can be built as follows: Let X be a measure space and Y a topological space, then $f : X \to Y$ is measurable iff $f^{-1}(B)$ is measurable for each Borel set B in Y. We don't use this abstraction in this text.

Definition 5.2.39. $f : X \to \mathbb{C}$ is said to be **integrable** if both $\operatorname{Re} f$ and $\operatorname{Im} f$ are integrable. In that case, we define the integral of f over X by

$$\int_X f\,d\mu = \int_X \operatorname{Re} f\,d\mu + i\int_X \operatorname{Im} f\,d\mu.$$

If $A \in \Sigma$, we define the integral of f over A by

$$\int_A f\,d\mu = \int_X \chi_A f\,d\mu.$$

From now on, in case if $f : X \to \mathbb{C}$ is integrable we extend the meaning of the symbol $\mathcal{L}^1(X,\mu)$ and say $f \in \mathcal{L}^1(X,\mu)$. Again, for clarity, we use the terms **extended real-valued integrable function**, **real-valued integrable function** and **complex integrable function** to distinguish members in $\mathcal{L}^1(\mu)$.

As a routine work, we list all basic properties of our newly defined integral.

Theorem 5.2.40. *Let $f,g : X \to \mathbb{C}$ and $\alpha,\beta \in \mathbb{C}$.*

(i) If $f,g \in \mathcal{L}^1(X)$, then $\alpha f + \beta g \in \mathcal{L}^1(X)$ and

$$\int_X (\alpha f + \beta g)\,d\mu = \alpha\int_X f\,d\mu + \beta\int_X g\,d\mu.$$

(ii) If f is measurable, then $f \in \mathcal{L}^1(X)$ iff $|f| \in \mathcal{L}^1(X)$. In that case,

$$\left|\int_X f\,d\mu\right| \le \int_X |f|\,d\mu.$$

(iii) If $f \in \mathcal{L}^1(X)$ and $A \in \Sigma$, then $\chi_A f \in \mathcal{L}^1(X)$. If $A = A_1 \sqcup A_2$, $A_1,A_2 \in \Sigma$, then

$$\int_A f\,d\mu = \int_{A_1} f\,d\mu + \int_{A_2} f\,d\mu.$$

(iv) If $f \in \mathcal{L}^1(X)$ and $\{A_n\}$ is a disjoint collection in Σ, then the series $\sum_{n=1}^\infty \int_{A_n} f\,d\mu$ converges absolutely and

$$\int_{\bigsqcup_{n=1}^\infty A_n} f\,d\mu = \sum_{n=1}^\infty \int_{A_n} f\,d\mu.$$

Proof. (i) The integrability is straightforward. We break down the proof of linearity into two parts. Firstly, we show that $\int_X \alpha f\,d\mu = \alpha\int_X f\,d\mu$. Secondly, we show that $\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$, both are simple computations.

(ii) Let f be integrable, then $|f| = \sqrt{|\operatorname{Re} f|^2 + |\operatorname{Im} f|^2} \leq |\operatorname{Re} f| + |\operatorname{Im} f|$[(7)] shows that $|f|$ is integrable. Conversely, assume $|f|$ is integrable, then $|\operatorname{Re} f|$, $|\operatorname{Im} f| \leq |f|$ shows that $\operatorname{Re} f$ and $\operatorname{Im} f$ are integrable, thus f is integrable by definition. Assume $f \in \mathcal{L}^1(X)$, we can find an $\alpha \in \mathbb{C}$ such that $|\alpha| = 1$ and $\left|\int_X f\,d\mu\right| = \alpha \int_X f\,d\mu = \int_X \operatorname{Re}(\alpha f)\,d\mu + i\int_X \operatorname{Im}(\alpha f)\,d\mu = \int_X \operatorname{Re}(\alpha f)\,d\mu \leq \int_X |\alpha f|\,d\mu = \int_X |f|\,d\mu$.

(iii) $\chi_A f \in \mathcal{L}^1(X)$ because $|\chi_A f| \leq |f|$, showing that $|\chi_A f|$ is integrable, so is $\chi_A f$. Since $\chi_{A_1 \sqcup A_2} = \chi_{A_1} + \chi_{A_2}$, the last equality holds.

(iv) Let $A = \bigsqcup_{k=1}^{\infty} A_k$, then

$$\left|\int_X \chi_A f\,d\mu - \int_X \chi_{\bigsqcup_{k=1}^n A_k} f\,d\mu\right| \leq \int_X |\chi_A - \chi_{\bigsqcup_{k=1}^n A_k}||f|\,d\mu.$$

Since $|\chi_A - \chi_{\bigsqcup_{k=1}^n A_k}||f| \leq |f|$ and $|\chi_A - \chi_{\bigsqcup_{k=1}^n A_k}||f| \to 0$ pointwise on X, by Lebesgue dominated convergence theorem,

$$\int_X \chi_A f\,d\mu = \lim_{n\to\infty} \int_{\bigsqcup_{k=1}^n A_k} f\,d\mu = \sum_{k=1}^{\infty} \int_{A_k} f\,d\mu.$$

Since $A = \bigsqcup_{k=1}^{\infty} A_k = \bigsqcup_{k=1}^{\infty} A_{\sigma(k)}$, for every bijection $\sigma : \mathbb{N} \to \mathbb{N}$, the value of the series $\sum_{k=1}^{\infty} \int_{A_k} f\,d\mu$ is independent of the rearrangement of the summands, hence converges absolutely. ☐

Note that again we have continuity of integration due to countable additivity, as remarked earlier. Moreover, let $f, g : X \to \mathbb{R}$ be real-valued integrable functions, by $|\int_X (f + ig)\,d\mu| \leq \int_X |f + ig|\,d\mu$, one has an interesting inequality:

$$\sqrt{\left(\int_X f\,d\mu\right)^2 + \left(\int_X g\,d\mu\right)^2} \leq \int_X \sqrt{f^2 + g^2}\,d\mu.$$

Theorem 5.2.41 (Lebesgue's Dominated Convergence, Complex Form). *Let $f, f_1, f_2, \cdots : X \to \mathbb{C}$ be measurable functions and $g : X \to [0,\infty]$ integrable such that $f_n \to f$ pointwise a.e. on X and*

$$|f_n| \leq g \text{ a.e. on } X,$$

then f is integrable over X and

$$\lim_{n\to\infty} \int_X f_n\,d\mu = \int_X f\,d\mu.$$

Proof. $f \in \mathcal{L}^1(X)$ as $|f| \leq g$ a.e. on X. Since

$$\left|\int_X f\,d\mu - \int_X f_n\,d\mu\right| \leq \int_X |f - f_n|\,d\mu,$$

and $|f - f_n| \leq 2g$ a.e. with $|f - f_n| \to 0$ pointwise a.e. on X, hence the result follows from Lebesgue dominated convergence theorem for extended real-valued measurable functions. ☐

Theorem 5.2.42. *Let $f \in \mathcal{L}^1(X, \mu)$.*

[(7)]For $x, y \geq 0$, $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$ iff $x + y \leq x + y + 2\sqrt{xy}$, the latter holds obviously.

(i) If $\int_E f\,d\mu = 0$ for every $E \in \Sigma$, then $f = 0$ a.e. on X.

(ii) If $|\int_X f\,d\mu| = \int_X |f|\,d\mu$, then there is a constant $\alpha \in \mathbb{C}$ such that $\alpha f = |f|$ a.e. on X.

Proof. (i) write $f = u + iv$, where u,v are real-valued, then $\int_E u\,d\mu = \int_E v\,d\mu = 0$ for every $E \in \Sigma$, and the rest is left as exercises.

(ii) Since there is $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ such that $\alpha \int_X f\,d\mu = |\int_X f\,d\mu|$, we have $\int_X (|f| - \alpha f)\,d\mu = 0$. Write $\alpha f = u + iv$, then $\int_X (|f| - u)\,d\mu = \int_X v\,d\mu = 0$. Since $|f| = |\alpha f| = |u + iv| \geq |u| \geq u$, so $|f| = u$ a.e. (on X). This also implies $\alpha f = |f| + iv$ a.e.. As $|\alpha| = 1$, $v = 0$ a.e., so $\alpha f = |f|$ a.e.. □

Theorem 5.2.43. *Suppose $\mu(X) < \infty$, $f \in \mathcal{L}^1(X,\mu)$, S is a closed set in $\mathbb{C}$ and*

$$\frac{1}{\mu(E)} \int_E f\,d\mu \in S$$

for every $E \in \Sigma$ with $\mu(E) > 0$, then $f(x) \in S$ for a.e. $x \in X$.

Proof. We may assume f is a complex measurable function. Since $\mathbb{C} - S$ is open, there are countably many open balls B_i such that $\mathbb{C} - S = \bigcup B_i$, hence $f^{-1}(\mathbb{C} - S) = \bigcup f^{-1}(B_i)$. Let $B_i = B(x_i, r_i) \subseteq \mathbb{C} - S$, it is enough to show $f^{-1}(B_i)$ has μ-measure zero for each i, suppose not, then

$$\left| \frac{1}{\mu(f^{-1}(B_i))} \int_{f^{-1}(B_i)} f\,d\mu - x_i \right| \leq \frac{1}{\mu(f^{-1}(B_i))} \int_{f^{-1}(B_i)} |f(x) - x_i|\,d\mu < r_i,$$

meaning that $S \cap B_i \neq \emptyset$, a contradiction. □

5.3 Product Measures

5.3.1 Definitions of Product Measure and Product sigma-algebra

As pointed out earlier the construction of Lebesgue measure can be abstracted to construct many more measures with the help of Carathéodory-Hahn Theorem 4.3.14. As an application, we have used this extension theorem to construct Lebesgue-Stieltjes measure on $\mathbb{R}$. With the help of this extension theorem again, we will develop an important fact that given two measure spaces X and Y one can always construct, in a reasonable manner, a new measure called product measure defined on a nice enough σ-algebra on $X \times Y$. Once this is done, we will also prove the Fubini-Tonelli theorem which enables us to switch the order of iterated integrals, an important technique even dealing with integrals which take the form $\int_a^b f\,dx$!

In this section we fix $X, Y, \mathcal{M}, \mathcal{N}, \mu, \nu$ to be the components of two measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$.

Definition 5.3.1. We define

$$\mathcal{M} \times \mathcal{N} := \{A \times B : A \in \mathcal{M}, B \in \mathcal{N}\}$$

to be the collection of **measurable rectangles** in $X \times Y$.

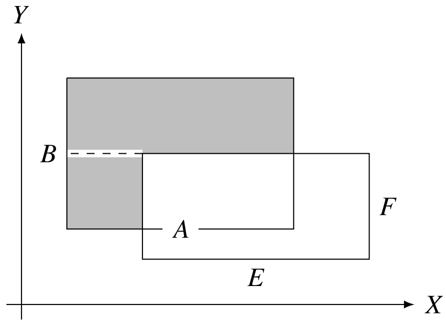


Figure 5.5: Subtraction of rectangles.

Let $A \times B, E \times F \in \mathcal{M} \times \mathcal{N}$, we have $(A \times B) \cap (E \times F) = (A \cap E) \times (B \cap F)$. By Figure 5.5 we have $A \times B - E \times F = \big(A \times (B - F)\big) \sqcup \big((A - E) \times (B \cap F)\big)$.

Hence $\mathcal{M} \times \mathcal{N}$ is closed under finite intersection and subtraction of measurable rectangles is a union of finite disjoint unions of measurable rectangles, and hence $\mathcal{M} \times \mathcal{N}$ forms a semiring.

To get a measure, we first introduce the following set functions on $\mathcal{M} \times \mathcal{N}$: If $A \times B \in \mathcal{M} \times \mathcal{N}$, define

$$\lambda(A \times B) = \mu(A)\nu(B). \tag{5.3.2}$$

This definition is natural in the sense that if μ and ν are Lebesgue measure, then λ is nothing but the area function defined on rectangles in $\mathbb{R}^2$. In order to extend it to a measure, we have to check that λ is a premeasure on $\mathcal{M} \times \mathcal{N}$.

Proposition 5.3.3. *$\lambda : \mathcal{M} \times \mathcal{N} \to [0, \infty]$ defined in (5.3.2) is a premeasure.*

Proof. By Proposition 4.3.22 we only need to show λ possesses countable additivity. Let $\{A_i \times B_i\}_{i=1}^{\infty}$ be a countable disjoint collection in $\mathcal{M} \times \mathcal{N}$ such that $A \times B = \bigsqcup_{i=1}^{\infty}(A_i \times B_i) \in \mathcal{M} \times \mathcal{N}$. Since

$$\chi_A(x)\chi_B(y) = \chi_{A \times B}(x, y) = \sum_{i=1}^{\infty} \chi_{A_i \times B_i}(x, y) = \sum_{i=1}^{\infty} \chi_{A_i}(x)\chi_{B_i}(y), \tag{5.3.4}$$

we integrate both sides of (5.3.4) and apply monotone convergence theorem to obtain

$$\begin{aligned} \mu(A)\chi_B(y) &= \int_X \sum_{i=1}^{\infty} \chi_{A_i}(x)\chi_{B_i}(y)\, d\mu(x) \\ &= \sum_{i=1}^{\infty} \int_X \chi_{A_i}(x)\chi_{B_i}(y)\, d\mu(x) = \sum_{i=1}^{\infty} \mu(A_i)\chi_{B_i}(y). \end{aligned} \tag{5.3.5}$$

Next we integrate both sides of (5.3.5) and apply monotone convergence theorem once more to obtain

$$\mu(A)\nu(B) = \sum_{i=1}^{\infty} \mu(A_i)\nu(B_i),$$

i.e., λ is countably additive, proving that λ is a premeasure. $\square$

By part (i) of Carathéodory-Hahn Theorem 4.3.14 we can now extend λ to a measure on the smallest σ-algebra containing all measurable rectangles. However, such extension may not be unique. In order to get a unique extension we should require the space $X \times Y$ be σ-finite (with respect to the length on $\mathcal{M} \times \mathcal{N}$).

In the sequel we shall require both X and Y be σ-finite (we will indicate this in each of the results). In that case, λ is σ-finite because there are $X_i \in \mathcal{M}$ and $Y_i \in \mathcal{N}$ such that $X = \bigcup X_i$ and $Y = \bigcup Y_j$ with $\mu(X_i), \nu(Y_j) < \infty$, it follows that

$$X \times Y = \bigcup_i \bigcup_j (X_i \times Y_j)$$

and $\lambda(X_i \times Y_j) = \mu(X_i)\nu(Y_j) < \infty$. Let the outer measure induced by λ be denoted by λ^*. For the moment we shall not deal with the σ-algebra of λ^*-measurable subsets on $X \times Y$, what we are going to do is to develop the theory on a "cleaner" σ-algebra defined by

$$\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{M} \times \mathcal{N}).$$

Namely, the σ-algebra generated by the measurable rectangles. We summarize them as a definition.

Definition 5.3.6. Let X and Y be measure spaces. We call $\mathcal{M} \otimes \mathcal{N}$ the **product σ-algebra**. If X and Y are also σ-finite, we denote

$$\mu \times \nu = \lambda^*|_{\mathcal{M} \otimes \mathcal{N}}$$

the unique measure on the measurable space $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ that extends λ, called **product measure**.

More generally, let (X_i, Σ_i, μ_i), $i = 1, 2, \ldots, n$, be measure spaces, it is easy to check $\Sigma_1 \times \cdots \times \Sigma_n := \{A_1 \times \cdots \times A_n : A_i \in \Sigma_i\}$ is a semiring and the set function $A_1 \times \cdots \times A_n \mapsto \mu_1(A_1) \cdots \mu_n(A_n)$ defined on it is a premeasure which can be extended to a measure on a σ-algebra containing $\bigotimes_{i=1}^n \Sigma_i := \Sigma_1 \otimes \cdots \otimes \Sigma_n := \sigma(\Sigma_1 \times \cdots \times \Sigma_n)$.

By letting $\mu_1, \ldots, \mu_n = m$, the Lebesgue measure on $\mathbb{R}$, then the completion of $m \times \cdots \times m$ is the n-dimensional Lebesgue measure on $\mathbb{R}^n$, which is of our great interest of course! For that purpose we will study the product of more than two σ-algebras latter. Let's for simplicity stick with the case $n = 2$.

5.3.2 Sections of Sets and Functions, Monotone Class Lemma

Definition 5.3.7. Let X, Y be two measure spaces. For $E \subseteq X \times Y$, we define the ***x*-section** E_x and ***y*-section** E^y of E by

$$E_x = \{y \in Y : (x, y) \in E\} \quad \text{and} \quad E^y = \{x \in X : (x, y) \in E\}.$$

Suppose that f is a function on $X \times Y$, we define the ***x*-section** f_x and ***y*-section** f^y of f by

$$f_x(y) = f^y(x) = f(x, y).$$

For example, consider the closed region $D \subseteq \mathbb{R}^2$ in Figure 5.6. Given $x \in \mathbb{R}$, the section D_x is those y such that $(x, y) \in D$, i.e., the projection of the dashed segment to the y-axis.

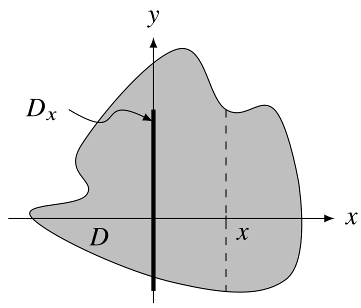


Figure 5.6: Example of x-section.

Proposition 5.3.8. *Let A be an index set and E, E_α, F, F_α be subsets of $X \times Y$, the following holds for all $x \in X$ and $y \in Y$.*

(i) $(\chi_E)_x(y) = \chi_{E_x}(y)$ *and* $(\chi_E)^y(x) = \chi_{E^y}(x)$.

(ii) $(\bigcup_{\alpha \in A} E_\alpha)_x = \bigcup_{\alpha \in A}(E_\alpha)_x$ *and* $(\bigcup_{\alpha \in A} E_\alpha)^y = \bigcup_{\alpha \in A}(E_\alpha)^y$.

(iii) $(\bigcap_{\alpha \in A} E_\alpha)_x = \bigcap_{\alpha \in A}(E_\alpha)_x$ *and* $(\bigcap_{\alpha \in A} E_\alpha)^y = \bigcap_{\alpha \in A}(E_\alpha)^y$.

(iv) $(E - F)_x = E_x - F_x$ *and* $(E - F)^y = E^y - F^y$.

(v) If $E \subseteq F$, *then* $E_x \subseteq F_x$ *and* $E^y \subseteq F^y$.

The proof is left as an exercise.

Proposition 5.3.9.

(i) If $E \in \mathcal{M} \otimes \mathcal{N}$, *then* $E_x \in \mathcal{N}$ *and* $E^y \in \mathcal{M}$, *for all* $x \in X, y \in Y$.

(ii) If f *is* $\mathcal{M} \otimes \mathcal{N}$*-measurable, then* f_x *is* $\mathcal{N}$*-measurable for all* $x \in X$ *and* f^y *is* $\mathcal{M}$*-measurbale for all* $y \in Y$.

Proof. (i) Let $x \in X, y \in Y$ be fixed. Consider the collection

$$\mathcal{C} := \{E \subseteq X \times Y : E_x \in \mathcal{N}, E^y \in \mathcal{M}\}.$$

Each measurable rectangle $A \times B \in \mathcal{M} \times \mathcal{N}$ is in $\mathcal{C}$ because $(A \times B)_x = B$ if $x \in A$ and $= \emptyset$ otherwise, similarly $(A \times B)^y \in \mathcal{M}$. It remains to show $\mathcal{C}$ is a σ-algebra, which follows easily from Proposition 5.3.8, hence (i) follows.

(ii) It follows immediately from (i) because for every set E, $f_x^{-1}(E) = (f^{-1}(E))_x$ and $(f^y)^{-1}(E) = (f^{-1}(E))^y$. □

Before we proceed, we need a technical lemma which provides a simple proof to Theorem 5.3.12 from which Theorem 5.3.14 almost directly follows. We need some terminology to begin with.

Definition 5.3.10. Let X be a space, a subset $\mathcal{C}$ of 2^X is said to be a **monotone class on *X*** provided it has the following properties:

(i) It is closed under **countable increasing union**:

$$E_i \in C, E_1 \subseteq E_2 \subseteq \cdots \implies \bigcup_{i=1}^{\infty} E_i \in C.$$

(ii) It is closed under **countable decreasing intersection**:

$$E_i \in C, E_1 \supseteq E_2 \supseteq \cdots \implies \bigcap_{i=1}^{\infty} E_i \in C.$$

It is clear that every σ-algebra is a monotone class. By direct verification arbitrary intersection of monotone classes is still a monotone class. So for each subset $\mathcal{A}$ of 2^X we can speak of the unique smallest monotone class that contains $\mathcal{A}$, denoted by $\boldsymbol{\mathcal{M}o(\mathcal{A})}$. It turns out that:

Lemma 5.3.11 (Monotone Class). *If $\mathcal{A}$ is an algebra of subsets of X, then*

$$\mathcal{M}o(\mathcal{A}) = \sigma(\mathcal{A}).$$

In particular, as remarked before if $\mathcal{S}$ is a semiring then $\mathcal{S}_\sqcup \cup \{X\}$ becomes an algebra, and hence $\mathcal{M}o(\mathcal{S}_\sqcup \cup \{X\}) = \sigma(\mathcal{S}_\sqcup \cup \{X\}) \supseteq \sigma(\mathcal{S})$.

Proof. Since $\sigma(\mathcal{A})$ is a σ-algebra, $\mathcal{M}o(\mathcal{A}) \subseteq \sigma(\mathcal{A})$. To show the reverse inclusion, it remains to prove that $\mathcal{M}o(\mathcal{A})$ is a σ-algebra. For $A \subseteq X$, define

$$C(A) = \{B \subseteq X : B - A, A - B, A \cup B \in \mathcal{M}o(\mathcal{A})\},$$

it is easy to check $C(A)$ is a monotone class and $B \in C(A)$ iff $A \in C(B)$. Now if $A \in \mathcal{A}$, then for each $B \in \mathcal{A}$, $B \in C(A)$ since $\mathcal{A}$ is an algebra, and hence $\mathcal{M}o(\mathcal{A}) \subseteq C(A)$. But this is also true for each $A \in \mathcal{A}$, hence **for every $\boldsymbol{B \in \mathcal{M}o(\mathcal{A})}$**, $B \in C(A)$, for all $A \in \mathcal{A}$, i.e., $A \in C(B)$ for all $A \in \mathcal{A}$, and hence $\mathcal{M}o(\mathcal{A}) \subseteq C(B)$, so **for every $\boldsymbol{A \in \mathcal{M}o(\mathcal{A})}$**, $A - B, A \cup B \in \mathcal{M}o(\mathcal{A})$. What's more, $X \in \mathcal{A}$, hence $X \in \mathcal{M}o(\mathcal{A})$.

It remains to show $\mathcal{M}o(\mathcal{A})$ is closed under countable union. Let $A_i \in \mathcal{M}o(\mathcal{A})$, $i = 1, 2, \ldots$, then $\bigcup_{i=1}^n A_i \in \mathcal{M}o(\mathcal{A})$ for each n, hence we are done. $\square$

5.3.3 Fubini-Tonelli Theorem

Clean Version

Theorem 5.3.12. *Let X, Y be σ-finite measure spaces. If $E \in \mathcal{M} \otimes \mathcal{N}$, then the functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable on X and Y respectively. Moreover,*

$$\mu \times \nu(E) = \int_X \nu(E_x)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y). \tag{5.3.13}$$

Proof. Let C be the collection of subsets of $X \times Y$ for which the theorem holds. Since for each $A \times B \in \mathcal{M} \times \mathcal{N}$, $\nu((A \times B)_x) = \chi_A(x)\nu(B)$ and $\mu((A \times B)^y) = \mu(A)\chi_B(y)$, so clearly $\mathcal{S} := \mathcal{M} \times \mathcal{N} \subseteq C$. By finite additivity of measures, $\mathcal{S}_\sqcup \subseteq C$, so it remains to show C is a monotone class.

We first assume μ and ν are finite measures. Let $E_1, E_2, \cdots \in C$ be ascending and write $E = \bigcup E_n$. By continuity of measure $\nu((E_n)_x) \nearrow \nu((E)_x)$ and $\nu((E_n)^y) \nearrow \nu((E)^y)$

pointwise on X and Y respectively, hence $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable and (5.3.13) holds by monotone convergence theorem. Thus C is closed under countable increasing union.

Similarly, let $E_1, E_2, \cdots \in C$ be descending and write $E = \bigcap E_n$. Since $\nu((E_1)_x)$, $\mu((E_1)^y) < \infty$, by continuity of measure both $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable and as $\nu((E_n)_x) \le \nu((E_1)_x)$ and $\mu((E_n)^y) \le \mu((E_1)^y)$ for each n, hence (5.3.13) holds by dominated convergence theorem and we conclude C is also closed under countable decreasing intersection, so C is a monotone class. Next, $C \supseteq \mathcal{M}o(\mathcal{S}_\sqcup) = \sigma(\mathcal{S}_\sqcup) \supseteq \sigma(\mathcal{S}) = \mathcal{M} \otimes \mathcal{N}$, hence the theorem is true when μ and ν are finite measures.

Finally when μ and ν are σ-finite. Let $E \in \mathcal{M} \otimes \mathcal{N}$ and let $X_i \in \mathcal{M}, Y_j \in \mathcal{N}$ be both ascending such that $\mu(X_i), \nu(Y_j) < \infty$ and $X = \bigcup_i X_i, Y = \bigcup_j Y_j$. Then $X_i \times Y_j$ is a finite measure subspace[8] and by the last paragraph,

$$\mu \times \nu(E \cap (X_i \times Y_j)) = \int_X \nu((E \cap (X_i \times Y_j))_x)\, d\mu(x) = \int_Y \mu((E \cap (X_i \times Y_j))^y)\, d\nu(y),$$

the theorem now follows from the continuity of measure and monotone convergence theorem, both twice. $\square$

Theorem 5.3.14 (Fubini-Tonelli, Incomplete Version). *Let X and Y be σ-finite measure spaces.*

(i) ***(Tonelli)*** *If $f : X \times Y \to [0,\infty]$ is $\mathcal{M} \otimes \mathcal{N}$-measurable, then the functions $x \mapsto \int_Y f_x\, d\nu$ and $y \mapsto \int_X f^y\, d\mu$ are $\mathcal{M} \otimes \mathcal{N}$-measurable. Moreover,*

$$\int_{X \times Y} f\, d(\mu \times \nu) = \int_X \left[\int_Y f_x\, d\nu(y) \right] d\mu(x) = \int_Y \left[\int_X f^y\, d\mu(x) \right] d\nu(y). \tag{5.3.15}$$

(ii) ***(Fubini)*** *If $f \in \mathcal{L}^1(\mu \times \nu)$, then $x \mapsto \int_Y f_x\, d\nu, y \mapsto \int_X f^y\, d\mu$ are also integrable, $f_x \in \mathcal{L}^1(\nu), f^y \in \mathcal{L}^1(\mu)$ a.e. and (5.3.15) holds.*

Proof. (i) If f is a characteristic function, then equation (5.3.15) reduces to (5.3.13), hence by linearity of integrals (5.3.15) holds for nonnegative simple functions. As f is nonnegative measurable, by simple approximation theorem there is an increasing sequence of $\mathcal{M} \otimes \mathcal{N}$-measurable nonnegative simple functions which converges to f pointwise on $X \times Y$, so measurability of the functions in the statement of the theorem follows from Theorem 5.3.12. Finally the equalities follow from monotone convergence theorem.

(ii) If f is integrable over $X \times Y$, then apply Tonelli theorem to positive part and negative part of $\operatorname{Re} f$ and $\operatorname{Im} f$ to conclude Fubini's theorem. $\square$

Remark. When writing an iterated integral, we usually omit the brackets and write (5.3.15) as

$$\int_X \left[\int_Y f_x\, d\mu(y) \right] d\nu(x) = \int_X \int_Y f(x,y)\, d\mu(y) d\nu(x) = \int_X \int_Y f\, d\mu d\nu.$$

We omit those x and y when it is understood in the content. Also without confusion some may also write $\int_X \int_Y$ and $\int_Y \int_X$ as simply $\iint$.

[8] Note that $(\mathcal{M} \cap X_i) \otimes (\mathcal{N} \cap Y_j) = (\mathcal{M} \otimes \mathcal{N}) \cap (X_i \times Y_j)$.

Unclean Version

It is worth noting that even μ and ν are complete, $\mu\times\nu$ is very rare to be a complete measure on $X\times Y$. To see this, let $A\in\mathcal{M}$ with $\mu(A)=0$ and $B\notin\mathcal{N}$, Proposition 5.3.9 tells us $A\times B\notin\mathcal{M}\otimes\mathcal{N}$. But $A\times B\subseteq A\times Y$ and $\mu\times\nu(A\times Y)=0$. Moreover, It is possible that a function is measurable with respect to $\overline{\mathcal{M}\otimes\mathcal{N}}$ but not $\mathcal{M}\otimes\mathcal{N}$. That prompts us to work with completion of a measure in an attempt to enlarge the class of measurable functions for which the order of iterated integrals of them can be switched.

Hence the statement of incomplete version of Fubini-Tonelli theorem needs to be reformulated, which we shall see in Theorem 5.3.18. To prove this, let's go through the following lemmas.

Lemma 5.3.16. *Let X and Y be complete σ-finite measure spaces.*

(i) If $E\in\mathcal{M}\otimes\mathcal{N}$ and $\mu\times\nu(E)=0$, then $\nu(E_x)=\mu(E^y)=0$ for a.e. x and a.e. y.

(ii) If f is $\overline{\mathcal{M}\otimes\mathcal{N}}$-measurable and $f=0$ $\overline{\mu\times\nu}$-a.e., then for a.e. x and y, $f_x=0$ a.e. and $f^y=0$ a.e..

Proof. (i) Let $f=\chi_E$ in Theorem 5.3.14, then $\int_X\int_Y f_x\,d\nu(y)d\mu(x)=0$ implies

$$\int_Y f_x\,d\nu(y)=\int_Y \chi_{E_x}(y)\,d\nu(y)=\nu(E_x)=0$$

for a.e. x. Similarly $\mu(E^y)=0$ for a.e. y.

(ii) Let $\{f\neq 0\}=\{(x,y)\in X\times Y: f(x,y)\neq 0\}$. By hypothesis $\overline{\mu\times\nu}\{f\neq 0\}=0$ and hence there is a $Z\in\mathcal{M}\otimes\mathcal{N}$ such that $\{f\neq 0\}\subseteq Z$ and $\mu\times\nu(Z)=0$. By part (i) for a.e. x and a.e. y, $\nu(Z_x)=\mu(Z^y)=0$, so by completeness for a.e. x and a.e. y,

$$\nu(\{f\neq 0\}_x)=\mu(\{f\neq 0\}^y)=0.$$

Since $\{f\neq 0\}_x=\{f_x\neq 0\}$ and $\{f\neq 0\}^y=\{f^y\neq 0\}$, so for a.e. x and a.e. y, $f_x=0$ a.e. and $f^y=0$ a.e.. ☐

Although the completion of any measure space (X,Σ,μ) can always be done, the σ-algebra $\overline{\Sigma}$ eventually becomes very complicated. The following result shows that every $\overline{\Sigma}$-measurable function can be thought of a Σ-measurable function if we don't worry about a set of measure zero.

Lemma 5.3.17. *Let (X,Σ,μ) be a measure space and $(X,\overline{\Sigma},\overline{\mu})$ its completion. If f is an $\overline{\Sigma}$-measurable function on X, then there is a Σ-measurable function g such that $f=g$ $\overline{\mu}$-a.e. on X.*

Proof. When $f=\chi_E$ for some $E\in\overline{\Sigma}$, then E takes the form $E=A\cup B$ where $A\in\Sigma$ and $\overline{\mu}(B)=0$. For sure we have $f=\chi_A$ $\overline{\mu}$-a.e. and χ_A is Σ-measurable.

For the general case, we apply simple approximation theorem to get a sequence of simple functions $\{\phi_n\}$ such that $\phi_n\to f$ pointwise on all of X. As in the first paragraph we can construct a Σ-measurable simple function ψ_n so that $\phi_n=\psi_n$ except on Z_n, where $\overline{\mu}(Z_n)=0$. Then $g:=\lim_{n\to\infty}\chi_{X-\bigcup_k Z_k}\psi_n$ is Σ-measurable and $f=g$ except possibly on $\bigcup Z_n$. ☐

Theorem 5.3.18 (Fubini-Tonelli, Complete Version). *Let X and Y be complete σ-finite measure spaces and $(X\times Y,\overline{\mathcal{M}\otimes\mathcal{N}},\overline{\mu\times\nu})$ the completion of $(X\times Y,\mathcal{M}\otimes$*

$\mathcal{N}, \mu\times\nu)$. Let f be $\overline{\mathcal{M}\otimes\mathcal{N}}$-measurable and consider case **(a)** *$f \geq 0$ and case* **(b)** *$f \in \mathcal{L}^1(\overline{\mu\times\nu})$.*

(i) In case (a) and (b), for a.e. x and y, f_x is $\mathcal{N}$-measurable and f^y is $\mathcal{M}$-measurable. Moreover, $x \mapsto \int_Y f_x\, d\nu$ and $y \mapsto \int_X f^y\, d\mu$ are measurable.

(ii) Furthermore, in case (a), one has

$$\int_{X\times Y} f\, d\overline{\mu\times\nu} = \int_Y\int_X f^y\, d\mu d\nu = \int_X\int_Y f_x\, d\nu d\mu,$$

the equality also holds in case (b).

Proof. (i) By Lemma 5.3.17 we can find an $\mathcal{M}\otimes\mathcal{N}$-measurable function g such that $f = g$ $\overline{\mu\times\nu}$-a.e.. By (ii) of Lemma 5.3.16 for a.e. x and y, $f_x - g_x = 0$ a.e. and $f^y - g^y = 0$ a.e.. Since g is $\mathcal{M}\otimes\mathcal{N}$-measurable, by Proposition 5.3.9 and by completeness of μ and ν we conclude for a.e. x and y, f_x is $\mathcal{N}$-measurable and f^y is $\mathcal{M}$-measurable. The a.e. defined maps

$$x \mapsto \int_Y f_x\, d\nu = \int_Y g_x\, d\nu \quad \text{and} \quad y \mapsto \int_X f^y\, d\mu = \int_X g^y\, d\mu$$

are both measurable by (i) of Fubini-Tonelli Theorem 5.3.14 and completeness of μ and ν.

(ii) Finally in case (a) and (b), since

$$\int_{X\times Y} f\, d\overline{\mu\times\nu} = \int_{X\times Y} g\, d\overline{\mu\times\nu} = \int_{X\times Y} g\, d(\mu\times\nu),$$

the result also follows from incomplete version of Fubini-Tonelli theorem. □

In application when we try to compute $\iint f\, d\mu d\nu$, we usually try to first compute $\iint |f|\, d\mu d\nu$ or $\iint |f|\, d\nu d\mu$, if one of them is finite, then by Fubini-Tonelli Theorem $\iint |f|\, d(\mu\times\nu)$ is also finite, hence f is integrable with respect to the product measure (or its completion), and hence by Fubini-Tonelli Theorem again we can interchange the order of integration. Hopefully in one of the orders the integration is easier to compute.

Example 5.3.19. We try to compute $\displaystyle\int_0^\infty \frac{\cos x - 1}{xe^x}\, dx$. Write

$$\begin{aligned}\int_0^\infty \frac{1-\cos y}{ye^y}\, dy &= 2\int_0^\infty \frac{1}{ye^y}\int_{t=0}^{t=1} d\left(\sin^2(ty/2)\right)\, dy = \int_0^\infty\int_0^1 e^{-y}\sin(ty)\, dt dy \\ &= \int_0^1\int_0^\infty e^{-y}\sin(ty)\, dy dt. \end{aligned}\tag{5.3.20}$$

Due to absolute Riemann integrability dy can be replaced by $dm(y)$ and dt can be replaced by $dm(t)$. For simplicity let's leave the notation unchanged.

(5.3.20) follows because $(x,y) \mapsto e^{-y}\sin(ty)$ is $\mathcal{L}\otimes\mathcal{L}$-measurable(9) and

$$\int_0^\infty\int_0^1 |e^{-y}\sin(ty)|\, dt dy \leq \int_0^\infty\int_0^1 e^{-y}\, dt dy = 1 < \infty.$$

(9)Because the function is continuous, the preimage of (a,∞) under this function must be open, and since $\mathcal{L}\otimes\mathcal{L} \supseteq \mathcal{B}_\mathbb{R}\otimes\mathcal{B}_\mathbb{R} = \mathcal{B}_{\mathbb{R}^2}$ (easily seen by considering open rectangles with vertices lying on $\mathbb{Q}^2$), the function is thus $\mathcal{L}\otimes\mathcal{L}$ measurable.

The inner integral in (5.3.20) can be computed by integration by parts twice, thus

$$\int_0^\infty \frac{1-\cos x}{xe^x}\,dx = \int_0^1 \frac{t}{t^2+1}\,dt = \frac{\ln 2}{2}. \qquad \square$$

5.4 Product of More Than two sigma-Algebras

We are now able to generalize the theory to product of n (≥ 3) measures. The idea is simple, given measure spaces (X_i, Σ_i, μ_i), we can construct $\mu_1 \times \mu_2$ on $\Sigma_1 \otimes \Sigma_2$ and next $\lambda_1 := (\mu_1 \times \mu_2) \times \mu_3$ on $(\Sigma_1 \otimes \Sigma_2) \otimes \Sigma_3$. But hang on! We can also construct $\mu_2 \times \mu_3$ first and then $\lambda_2 := \mu_1 \times (\mu_2 \times \mu_3)$ on $\Sigma_1 \otimes (\Sigma_2 \otimes \Sigma_3)$. The first question is: Do we have

$$(\Sigma_1 \otimes \Sigma_2) \otimes \Sigma_3 = \Sigma_1 \otimes (\Sigma_2 \otimes \Sigma_3)?$$

If so, we then ask: Are the measures the same? The answer of first question is positive and we leave the proof that both σ-algebras are $\sigma(\Sigma_1 \times \Sigma_2 \times \Sigma_3)$ as an exercise. It is clear to us if X_i's are all σ-finite, then since both λ_1, λ_2 extend the set function on "measurable cubes": $A \times B \times C \mapsto \mu_1(A)\mu_2(B)\mu_3(C)$, they are indeed the same by Corollary 4.3.16.

In this section we aim at giving a more explicit description of these σ-algebras. To this end, we generalize the notion of product σ-algebras as follows:

Definition 5.4.1. Let $\{X_\alpha\}$ be an indexed collection of nonempty sets and define $\pi_\alpha : \prod_{\alpha \in A} X_\alpha \to X_\alpha$ the coordinate maps. Let Σ_α be the σ-algebra on X_α, we define the **product σ-algebra** on $\prod_{\alpha \in A} X_\alpha$ by

$$\bigotimes_{\alpha \in A} \Sigma_\alpha := \sigma\left\{\pi_\alpha^{-1}(S_\alpha) : S_\alpha \in \Sigma_\alpha, \alpha \in A\right\}.$$

In particular, if $A = \{1, 2, \ldots, n\}$, we write $\bigotimes_{\alpha \in A} \Sigma_\alpha = \bigotimes_{i=1}^n \Sigma_i$. If further $\Sigma_1 = \Sigma_2 = \cdots = \Sigma_n = \Sigma$, we write $\bigotimes_{i=1}^n \Sigma_i = \Sigma^{\otimes n}$.

Proposition 5.4.2. *If A is countable, then*

$$\bigotimes_{\alpha \in A} \Sigma_\alpha = \sigma\left\{\prod_{\alpha \in A} S_\alpha : S_\alpha \in \Sigma_\alpha\right\}.$$

Proof. If $S_\alpha \in \Sigma_\alpha$, then $\pi_\alpha^{-1}(S_\alpha) \in \bigotimes_{\alpha \in A} \Sigma_\alpha$ by definition, and hence $\prod_{\alpha \in A} S_\alpha = \bigcap_{\alpha \in A} \pi_\alpha^{-1}(S_\alpha) \in \bigotimes_{\alpha \in A} \Sigma_\alpha$. Conversely, it is obvious that for each $S_\alpha \in \Sigma_\alpha$, $\pi_\alpha^{-1}(S_\alpha)$ is contained in the RHS. $\square$

Proposition 5.4.3. *Suppose that $\Sigma_\alpha = \sigma(\mathcal{E}_\alpha)$, $\alpha \in A$.*

(i) $\bigotimes_{\alpha \in A} \Sigma_\alpha = \sigma\{\pi_\alpha^{-1}(E_\alpha) : E_\alpha \in \mathcal{E}_\alpha, \alpha \in A\}$.

(ii) If A is countable and $X_\alpha \in \mathcal{E}_\alpha$, then

$$\bigotimes_{\alpha \in A} \Sigma_\alpha = \sigma\left\{\prod_{\alpha \in A} E_\alpha : E_\alpha \in \mathcal{E}_\alpha\right\}.$$

Proof. (i) It is obvious that RHS is contained in LHS. To show the reverse inclusion, observe that for each $\alpha \in A$, $\{E \subseteq X_\alpha : \pi_\alpha^{-1}(E) \in \text{RHS}\}$ is a σ-algebra that contains $\mathcal{E}_\alpha$, and hence contains Σ_α, i.e., $\pi_\alpha^{-1}(S_\alpha) \in \text{RHS}$ for each $S_\alpha \in \Sigma_\alpha$. This is true for each α, and hence LHS $\subseteq$ RHS.

(ii) This follows from (i) as in the proof of Proposition 5.4.2. □

Proposition 5.4.4. *Let $(X_1,d_1),\ldots,(X_n,d_n)$ be metric spaces and let $X = \prod_{i=1}^n X_i$ be equipped with product metric defined by*

$$d((x_1,\ldots,x_n),(y_1,\ldots,y_n)) = \max_{1\le i\le n} d_i(x_i,y_i),$$

we have:

(i) $\bigotimes_{i=1}^n \mathcal{B}_{X_i} \subseteq \mathcal{B}_X$.

(ii) If X_i's are separable, then $\bigotimes_{i=1}^n \mathcal{B}_{X_i} = \mathcal{B}_X$.

Proof. (i) Since the collection of Borel sets on X_i is generated by the topology on X_i, hence by Proposition 5.4.3, $\bigotimes_{i=1}^n \mathcal{B}_{X_i}$ is generated by elements of the form $\prod_{i=1}^n U_i$, where U_i is open in X_i, and these elements are open in X, hence they are contained in $\mathcal{B}_X$.

(ii) Let C_i be a countable dense subset of X_i. Then $C_1 \times\cdots\times C_n$ is a countable dense subset of X and each open set in X can be expressed as a union of balls of the form $B_d((x_1,\ldots,x_n),r) = \prod_{i=1}^n B_{d_i}(x_i,r)$, where $x_i \in C_i$ and $r \in \mathbb{Q}$. As there are at most countably many such balls, hence each open set in X is contained in $\bigotimes_{i=1}^n \mathcal{B}_{X_i}$. □

Corollary 5.4.5. $\mathcal{B}_{\mathbb{R}}^{\otimes n} := \mathcal{B}_{\mathbb{R}} \otimes\cdots\otimes \mathcal{B}_{\mathbb{R}} = \mathcal{B}_{\mathbb{R}^n}$.

Proof. It is an immediate consequence of Proposition 5.4.4. □

Proposition 5.4.6. *Let (X_i,Σ_i,μ_i), $i = 1,2,\ldots,n$, be **σ-finite** measure spaces and let $1 \le m < n$, show that the product measure spaces*

$$\left(\left(\prod_{i=1}^m X_i \times \prod_{i=m+1}^n X_i\right), \left(\bigotimes_{i=1}^m \Sigma_i \otimes \bigotimes_{i=m+1}^n \Sigma_i\right), \left(\prod_{i=1}^m \mu_i \times \prod_{i=m+1}^n \mu_i\right)\right)$$

are the same as $(\prod_{i=1}^n X_i, \bigotimes_{i=1}^n \Sigma_i, \prod_{i=1}^n \mu_i)$.

Proof. It is left as an exercise. □

5.5 Lebesgue Measure on $\mathbb{R}^n$

We have noticed that the Lebesgue measure space $(\mathbb{R},\mathcal{L},m)$ on $\mathbb{R}$ is just the completion of $(\mathbb{R},\mathcal{B}_{\mathbb{R}},m)$[10]. It is reasonable to define Lebesgue measure on $\mathbb{R}^n$ as in the next definition.

Definition 5.5.1. Let m be the Lebesgue measure on $\mathbb{R}$ and $(\mathbb{R}^n,\mathcal{L}_n,m_n)$ the completion of $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}} \otimes\cdots\otimes \mathcal{B}_{\mathbb{R}}, m\times\cdots\times m)$, then $\mathcal{L}_n$ is the **Lebesgue σ-algebra on $\mathbb{R}^n$** and m_n is the **Lebesgue measure on $\mathbb{R}^n$**.

[10]Since each Lebesgue measurable set is a union of a F_σ set and a set of measure zero (Theorem 2.6.1).

Note that we can equivalently define Lebesgue measure space to be the completion of $(\mathbb{R}^n, \mathcal{L}\otimes\cdots\otimes\mathcal{L}, m\times\cdots\times m)$ (why?). Also by Corollary 5.4.5 we have $\mathcal{B}_{\mathbb{R}}^{\otimes n} = \mathcal{B}_{\mathbb{R}^n}$, hence each Borel set in $\mathbb{R}^n$ is $\mathcal{L}_n$-measurable.

In what follows if we say $E \subseteq \mathbb{R}^n$ is a rectangle, we mean $E = \prod_{i=1}^n E_i$, with E_i's $\subseteq \mathbb{R}$ called the sides of E. Further we say E is a measurable rectangle if each of its sides is Lebesgue measurable. We also remove the subscript letter n in m_n when there is no confusion.

Definition 5.5.2. Let m_1 be Lebesgue measure on $\mathbb{R}$, we denote $m^* : 2^{\mathbb{R}^n} \to [0,\infty]$ the outer measure induced by $\prod_{i=1}^n A_i \mapsto \prod_{i=1}^n m_1(A_i)$, $A_i \in \mathcal{L}$.

Let $(X_i, \Sigma_i, \lambda_i)$, $i = 1,2,\ldots,n$, be measure spaces, for $A_i \in \Sigma_i$, let $\lambda : \prod_{i=1}^n A_i \mapsto \prod_{i=1}^n \mu_i(A_i)$. Apart from doing completion, the space $X := X_1\times\cdots\times X_n$ equipped with the outer measure λ^* induced by λ and a σ-algebra, Σ^*, of λ^*-measurable sets is also a complete measure space that extends $\mathcal{S}$ and λ. It turns out that by Proposition 5.5.3 if X_i's are all σ-finite, then

$$(X, \Sigma^*, \lambda^*|_{\Sigma^*}) = (X, \overline{\bigotimes_{i=1}^n \Sigma_i}, \overline{\mu_1\times\cdots\times\mu_n}).$$

Proposition 5.5.3. *Let $\mathcal{S}$ be a semiring on a space X, $\lambda : \mathcal{S} \to [0,\infty]$ a pre-measure on $\mathcal{S}$ and λ^* an outer measure induced by λ. Denote Σ^* the collection of λ^*-measurable subsets of X. If λ is σ-finite, then $(X, \Sigma^*, \lambda^*|_{\Sigma^*})$ is the completion of $(X, \sigma(\mathcal{S}), \lambda^*|_{\sigma(\mathcal{S})})$.*

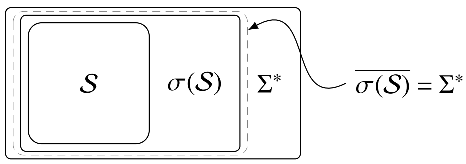


Figure 5.7: Completion of $\sigma(\mathcal{S})$ is Σ^*.

Recall that by completion $(X, \overline{\mathcal{M}}, \overline{\mu})$ of $(X, \mathcal{M}, \mu)$ we mean the unique measure space that is minimal in the sense that any other complete measure spaces extending $(X, \mathcal{M}, \mu)$ must also extend $(X, \overline{\mathcal{M}}, \overline{\mu})$.

Proof. For simplicity, let's write $\lambda^*|_{\Sigma^*} = \lambda^*$. It is obvious that $\Sigma^* \supseteq \overline{\sigma(\mathcal{S})}$ because $\Sigma^* \supseteq \sigma(\mathcal{S})$ and is itself complete. It remains to show $\Sigma^* \subseteq \overline{\sigma(\mathcal{S})}$. Let $E \in \Sigma^*$, as λ is σ-finite, there are $X_i \in \mathcal{S}$ such that $X = \bigcup X_i$ and $\lambda(X_i) < \infty$, and we have $E = \bigcup (E \cap X_i)$. Define $E_i' = E \cap X_i$ for simplicity. By Proposition 4.3.6 when a set has finite outer measure, we can approximate it by a $\mathcal{S}_{\sigma\delta}$ set from outside. In order to argue E_i' lies in $\overline{\sigma(\mathcal{S})}$, we prefer doing inner approximation. To this end, we approximate $X_i - E_i'$ from outside first.

Since $\lambda^*(X_i - E_i') < \infty$, there is a $A_i \in \mathcal{S}_{\sigma\delta}$ such that $A_i \supseteq X_i - E_i'$ and $\lambda^*(A_i) = \lambda^*(X_i - E_i')$. But then $\lambda^*(A_i - (X_i - E_i')) = 0$. Furthermore, $X_i - A_i \subseteq E_i'$, we expect $X_i - A_i \in \sigma(\mathcal{S})$ is a "good" inner approximation. Define $H_i = E_i' - (X_i - A_i)$ and write $E_i' = (X_i - A_i) \sqcup H_i$, we have

$$H_i = (E_i' - X_i) \cup (E_i' \cap A_i) = E_i' \cap A_i \subseteq A_i - (X_i - E_i').$$

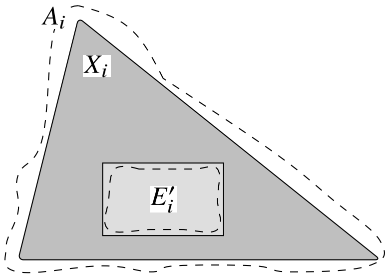


Figure 5.8: Approximate $X_i - E_i'$ from outside.

The rightmost one has λ^*-measure zero, but that means there is a $V_i \in \sigma(\mathcal{S})$ such that $V_i \supseteq H_i$ and $\lambda^*|_{\sigma(\mathcal{S})}(V_i) = 0$. We conclude $E_i' \in \overline{\sigma(\mathcal{S})}$, and hence $E = \bigcup E_i' \in \overline{\sigma(\mathcal{S})}$. □

In the case that all X_i's are Lebesgue measure space $\mathbb{R}$, the hypothesis in Proposition 5.5.3 is satisfied, and we have

$$(\mathbb{R}^n, \Sigma^*, m^*|_{\Sigma^*}) = (\mathbb{R}^n, \mathcal{L}_n, m).$$

As an immediate consequence:

Proposition 5.5.4. *A set $E \subseteq \mathbb{R}^n$ is Lebesgue measurable (i.e., being in $\mathcal{L}_n$) if and only if for any set $X \subseteq \mathbb{R}^n$,*

$$m^*(X) = m^*(X \cap E) + m^*(X - E). \tag{5.5.5}$$

Having defined m on $\mathbb{R}^n$, we deduce some of its basic properties. Specifically, we will show that m enjoys some regularity and every measurable set is "almost" a G_δ and F_σ set. We have seen their importance from m_1.

Proposition 5.5.6. *Let $E \in \mathcal{L}_n$.*

(i) $m(E) = \inf\{m(U) : U \supseteq E,\ U \text{ open}\} = \sup\{m(K) : K \subseteq E,\ K \text{ compact}\}$.

(ii) $E = F \cup N_1 = G - N_2$, *where F is F_σ, G is G_δ and $m(N_1) = m(N_2) = 0$.*

(iii) If $m(E) < \infty$, then for any $\epsilon > 0$, there is a finite collection $\{R_i\}_{i=1}^n$ of disjoint rectangles whose sides are intervals such that

$$m(E \Delta \bigsqcup_{i=1}^N R_i) < \epsilon.$$

Proof. (i) Since $m(E) = m^*(E)$, hence given $\epsilon > 0$, we can find nonempty measurable rectangles $\{R_i\}_{i=1}^\infty$ such that $\bigcup R_i \supseteq E$ and $\sum m(R_i) \le m(E) + \epsilon$. For each i, since $R_i = \prod_{i=j}^n A_{ij}$, $A_{ij} \in \mathcal{L}$, by approximating each A_{ij} from outside by an open set, we can find a rectangle $U_i \supseteq R_i$ whose sides are open set in $\mathbb{R}$ such that $m(U_i) \le m(R_i) + \epsilon/2^i$, hence

$$m\left(\bigcup U_i\right) \le \sum m(U_i) \le \sum m(R_i) + \epsilon \le m(E) + 2\epsilon. \tag{5.5.7}$$

As $\bigcup U_i$ is open, we are done. To prove inner regularity, we can imitate the proof of Theorem 4.3.29.

(ii) The proof is the case $n = 1$, we leave it as an exercise.

(iii) We start with (5.5.7) with U_i defined same as before, i.e., $U_i = \prod_{j=1}^n O_{ij}$ and has finite measure, where O_{ij} is open in $\mathbb{R}$ and hence can be expressed as a disjoint union of bounded open intervals. We can find a rectangle $V_i \subseteq U_i$ whose jth side is a union of finitely many intervals in O_{ij} such that $m(U_i) - m(V_i) < \epsilon/2^i$. Define $V = \bigcup_{i=1}^N V_i$, then V is a disjoint union of rectangles satisfying for each N,

$$m(E-V) \subseteq m\left(\bigcup_{i=1}^{\infty} U_i - V\right) \leq m\left(\bigcup_{i=1}^{N}(U_i - V_i)\right) + m\left(\bigcup_{i=N+1}^{\infty} U_i\right),$$

and

$$m(V-E) \leq m\left(\bigcup_{i=1}^{\infty} U_i - E\right) = m\left(\bigcup_{i=1}^{\infty} U_i\right) - m(E) < 2\epsilon.$$

Since $m(\bigcup_{i=1}^N (U_i - V_i)) \leq \sum_{i=1}^N (m(U_i) - m(V_i)) < \epsilon$, by choosing N large enough and ϵ small enough at the beginning, we are done. $\square$

Remark. In the proof of (i) of Proposition 5.5.6 we have actually shown that for any $E \subseteq \mathbb{R}^n$,

$$m^*(E) = \inf\{m(U) : U \supseteq E,\, U \text{ open}\}.$$

Remark. By part (iii) of Proposition 5.5.6 we are able to show the collection of compactly supported continuous functions $C_c(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n, m)$, as in the case on $\mathbb{R}$. This will be in Problem 5.41 for interested readers.

When doing Lebesgue measure on $\mathbb{R}$, certain results that are true for open sets can be almost immediately translated to measurable ones by outer regularity. The main property of open set in $\mathbb{R}$ that we use is: All of them are disjoint union of open intervals. The story is almost the same in $\mathbb{R}^n$, except that open sets are not necessarily disjoint union of open balls any more.

However, from the experience of multivariable Riemann integration when we try to argue a set "has volume" (or Jordan measurable), we try to do its inner approximation by filling rectangles contained in it as in Figure 5.9 with the "diameter" of the squares being smaller and smaller. It turns out that the same approach can be used to fill up all open sets in $\mathbb{R}^n$.

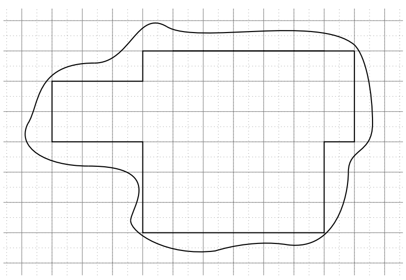


Figure 5.9: A step to construct inner Jordan measure.

In what follows, a **cube** is a rectangle whose sides are closed intervals of equal length.

Lemma 5.5.8. *Let $\mathcal{Q}_k$ be the collection of cubes whose sides have length 2^{-k} with vertices lying on the lattice $(2^{-k}\mathbb{Z})^n$. For $E \subseteq \mathbb{R}^n$, define*

$$J_-(E,k) = \bigcup_{Q\in\mathcal{Q}_k, Q\subseteq E} Q,$$

then $J_-(E,k)$ is ascending. If $E = U$ is open, then $U = \bigcup_{k=1}^{\infty} J_-(U,k)$ and U is a countable union of cubes with disjoint interiors.

Proof. Let $x \in J_-(E,k)$, then x is contained in a cube of length 2^{-k} contained in E. By dividing each side of the cube into half, $x \in J_-(E,k+1)$.

Assume $E = U$ is open, by definition $\bigcup_{k=1}^{\infty} J_-(U,k) \subseteq U$. Let $x \in U$, for each N there is $Q_N \in \mathcal{Q}_N$ such that $x \in Q_N$. If $y \in Q_N$, then $\|y-x\|_2 \le 2^{-N}\sqrt{n}$, hence we have $Q_N \subseteq B(x, 2^{-N}\sqrt{n})$. In particular, when N is large enough $x \in Q_N \subseteq U$, so $Q_N \subseteq J_-(U,N)$ and thus $x \in \bigcup_{k=1}^{\infty} J_-(U,k)$.

Finally U is a union of cubes with disjoint interiors by writing $U = J_-(U,1) \sqcup \bigsqcup_{k=2}^{\infty}(J_-(U,k) - J_-(U,k-1))$. ☐

Remark. Combining remark following Proposition 5.5.6 and Lemma 5.5.8, we have for any $E \subseteq \mathbb{R}^n$,

$$m^*(E) = \inf\left\{\sum m(I_i) : \bigcup I_i \supseteq E,\ I_i\text{'s are cubes}\right\}. \tag{5.5.9}$$

Some authors introduce the theory of Lebesgue measure on $\mathbb{R}^n$ by defining Lebesgue outer measure as in (5.5.9) and declare measurable subsets to be those satisfying (5.5.5). For different kinds of purposes, it is useful to keep these equivalent formulations in mind.

Definition 5.5.10. A measure defined on the σ-algebra of all Borel sets in a space X is called a **Borel measure on X**.

Theorem 5.5.11. *The Lebesgue measure on $\mathbb{R}^n$ has the following properties:*

(i) Let $E \in \mathcal{B}_{\mathbb{R}^n}$ and $x \in \mathbb{R}^n$, then $x + E \in \mathcal{B}_{\mathbb{R}^n}$ and

$$m(E) = m(x+E).$$

(ii) For every nonnegative Borel measurable function f on $\mathbb{R}^n$ and $y \in \mathbb{R}^n$,

$$\int_{\mathbb{R}^n} f(x+y)\,dm(x) = \int_{\mathbb{R}^n} f(x)\,dm(x) = \int_{\mathbb{R}^n} f(-x)\,dm(x).$$

(iii) Let μ be any σ-finite measure on $\mathcal{B}_{\mathbb{R}^n}$ such that $\mu(E+x) = \mu(E)$, for every $E \in \mathcal{B}_{\mathbb{R}^n}$ and $x \in \mathbb{R}^n$. Suppose

$$0 < \mu(E_0) = Cm(E_0) < \infty$$

for some $E_0 \in \mathcal{B}_{\mathbb{R}^n}$ and for some $C \ge 0$, then $\mu = Cm$.

Proof. (i) Let $E \in \mathcal{B}_{\mathbb{R}^n}$, consider the set $\mathcal{A} := \{E \subseteq \mathbb{R}^n : x + E \in \mathcal{B}_{\mathbb{R}^n}\}$. $\mathcal{A}$ contains all open sets as the map $y \mapsto x + y$ is a homeomorphism. By the following set equalities

$$x + (A - B) = (x+A) - (x+B) \quad \text{and} \quad x + \bigcup A_i = \bigcup (x + A_i),$$

it is easy to check $\mathcal{A}$ is a σ-algebra, hence $\mathcal{A} \supseteq \mathcal{B}_{\mathbb{R}^n}$, so $x + E \in \mathcal{B}_{\mathbb{R}^n}$.

Next to show $m(E) = m(x+E)$, we make use of the outer regularity of m. For each $\epsilon > 0$ we can find an open set $U \supseteq E$ such that $m(U) \le m(E) + \epsilon$. Now by Lemma 5.5.8 one can find cubes $Q_1, Q_2, \dots$ with disjoint interiors such that $\bigcup Q_i = U$, and it is an easy computation to verify $m(x + Q_i) = m(Q_i)$, hence $m(x + \bigcup_{i=1}^n Q_i) = m(\bigcup_{i=1}^n Q_i)$[(11)] and this implies $m(x+U) = m(U)$. As $x + U \supseteq x + E$, so

$$m(x+E) \le m(x+U) = m(U) \le m(E) + \epsilon,$$

for all $\epsilon > 0$, hence $m(x+E) \le m(E)$, but $m(E) \le m(x+E)$ directly follows.

(ii) It suffices to check it for $f = \chi_E$, where $E \in \mathcal{B}_{\mathbb{R}^n}$. The second equality may require a bit more work, but the technique used in (i) will do.

(iii) The statement $\mu = Cm$ is the same as $\mu(E) = Cm(E)$ for each Borel set E, which is the same as

$$m(E_0)\mu(E) = \mu(E_0)m(E)$$

for each Borel set E, this holds because

$$\begin{aligned} m(E_0)\mu(E) &= \int_{\mathbb{R}^n} m(E_0)\chi_E(y)\,d\mu(y) = \int_{\mathbb{R}^n} m(E_0 - y)\chi_E(y)\,d\mu(y) \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \chi_{E_0}(x+y)\chi_E(y)\,dm(x)d\mu(y), \end{aligned}$$

by incomplete version of Fubini-Tonelli theorem,

$$\begin{aligned} m(E_0)\mu(E) &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \chi_{E_0}(x+y)\chi_E(y)\,d\mu(y)dm(x) \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \chi_{E_0}(y)\chi_E(y-x)\,d\mu(y)dm(x) \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \chi_{E_0}(y)\chi_E(y-x)\,dm(x)d\mu(y) = \mu(E_0)m(E). \qquad \square \end{aligned}$$

(iii) of Theorem 5.5.11 is a very interesting result. Firstly, any translation invariant σ-finite Borel measure must be a constant multiple of Lebesgue measure. Secondly, any such Borel measure must be regular which we cannot tell directly by looking at translation invariant property. In fact there is a general theory stating that any finite Borel measure on a complete separable metric space must be regular (see Lemma 6.3.14), this generalizes the case that $X = \mathbb{R}^n$ by some σ-finite argument.

Lemma 5.5.12. *For $a \in \mathbb{R}^n$ and $r > 0$, $B(a,r) := \{x \in \mathbb{R}^n : \|x - a\|_2 < r\}$ satisfies*

$$m(B(0,r)) = r^n m(B(0,1)). \tag{5.5.13}$$

Proof. The equality $m(rE) = r^n m(E)$ holds obviously when E is a cube. By Lemma 5.5.8 for every open set U there are cubes $Q_1, Q_2, \dots$ having disjoint interiors such that $U = \bigcup Q_i$, by continuity of measure one has $m(rU) = r^n m(U)$. In particular, since $B(0,r) = rB(0,1)$, then (5.5.13) can be obtained by setting $U = B(0,1)$. $\square$

By Lemma 5.5.12 we can extend the result in Problem 2.6 if we insist on using 2-norm.

[(11)] Due to disjoint interior (i.e., pairwise intersection must be of measure zero).

Proposition 5.5.14. *Let $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ be a function that satisfies $\|f(x)-f(y)\|_2 \leq M\|x-y\|_2$, for some constant $M \geq 0$. Then for every set $A \subseteq \mathbb{R}^n$,*

$$m^*(\Phi(A)) \leq M^n m^*(A).$$

In particular, Φ takes Lebesgue measurable subsets to Lebesgue measurable subsets.

Proof. Since the proof is similar to the case $n = 1$, it is left as an exercise. □

Remark. The domain of Φ in Proposition 5.5.14 can be replaced by any subset X of $\mathbb{R}^n$ because any Lipschitz function on X can be extended to a Lipschitz function on $\mathbb{R}^n$. We give the detail in Problem 5.43. The case that $X = \mathbb{R}^n$ is good enough for most of the purpose in this section because all $n \times n$ (real) matrices are bounded linear transform, in particular, they are Lipschitz functions on $\mathbb{R}^n$.

5.6 Linear Algebra and Differentiation of Multivariable Functions

Let's recall some definitions and basic results in linear algebra and multivariable differentiation on Euclidean spaces. Knowledge in this section will be used in Section 5.7.2.

5.6.1 Basic Results and Operations of Differentiation

We will adopt the following convention: We always denote the coordinate function of $F : \mathbb{R}^n \to \mathbb{R}^m$ by f_i, i.e., $F = (f_1, f_2, \ldots, f_m)$. For a matrix $A : \mathbb{R}^n \to \mathbb{R}^m$ the p-norm $(1 \leq p \leq \infty)$ of A is denoted by

$$\|A\|_p = \sup\{\|Ax\|_p : x \in \mathbb{R}^n, \|x\|_p = 1\}.$$

$\|\cdot\|_p$ on $\mathbb{R}^{m\times n}$ is called matrix norm or operator norm. Special choices of p do have specific geometrical meanings, see Section **??**, for example. In addition to knowing definition of matrix norms, we are able to compute them explicitly in some cases. For example, let a_i's be column vectors of $\mathbb{K}^m$, from definition it is easy to show that

$$A = \left[a_1|\cdots|a_n\right] \implies \|A\|_1 = \max_{1\leq j\leq n} \|a_j\|_1 \tag{5.6.1}$$

$$A = \begin{bmatrix} a_1^t \\ \hline \vdots \\ \hline a_n^t \end{bmatrix} \implies \|A\|_\infty = \max_{1\leq i\leq m} \|a_i^t\|_1 \tag{5.6.2}$$

In words, $\|A\|_1$ is the maximum (absolute) column sum, while $\|A\|_\infty$ is the maximum (absolute) row sum.

Definition 5.6.3. If $F = (f_1, \ldots, f_m)$ is defined near $a \in \mathbb{R}^n$ and there is a linear transform $T : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\|x-a\|\to 0} \frac{\|F(x)-F(a)-T(x-a)\|}{\|x-a\|} = 0, \tag{5.6.4}$$

then we say f is **differentiable at a**. Furthermore, such linear transform is unique[12] and denoted by $DF(a)$. The matrix of $DF(a)$ with respect to the usual basis is denoted by $F'(a)$ or $JF(a)$. We say that **$F'(a)$ exists** if all its partial derivatives at a exist.

Remark. Recall that all norms on a finite dimensional vector space are *equivalent*, the differentiability of F is independent of any choice of norm. In particular, if we define $\|\cdot\| = \|\cdot\|_\infty$ in the numerator of (5.6.4), then F is differentiable at a iff all its coordinate functions are differentiable at a. From that we also conclude F is differentiable at a implies F is continuous at a.

Remark. When F is differentiable at a, the matrix of T satisfying (5.6.4) with respect to the usual basis is uniquely determined (called the **Jacobian matrix of F at a**) and can be computed explicitly. To see this, let T be a linear transform that satisfies (5.6.4), let $A \in \mathbb{R}^{m\times n}$ be its matrix with respect to usual basis $\{e_1, e_2, \ldots, e_n\}$. Choose $\|\cdot\| = \|\cdot\|_2$ in the domain, one has[13] $\lim_{x\to a} |f_i(x) - f_i(a) - e_i^t A(x-a)|/\|x-a\|_2 = 0$. Let $h \in \mathbb{R}$ and $h \to 0$, then $x = x(h) := a + he_j \to a$ and

$$0 = \lim_{h\to 0}\left|\frac{f_i(a+he_j) - f_i(a) - e_i^t A(he_j)}{h}\right| = \lim_{h\to 0}\left|\frac{f_i(a+he_j) - f_i(a)}{h} - e_i^t A e_j\right|,$$

hence $e_i^t A e_j = \frac{\partial f_i}{\partial x_j}(a)$ and $A = (\frac{\partial f_i}{\partial x_j}(a))_{i,j}$, which also shows that T is uniquely determined so that it is unambiguous to write $T = DF(a)$.

Remark. We may also write $F'(a) = (\frac{\partial f_i}{\partial x_j}(a))_{i,j}$ as $F_x(a)$ or $\frac{\partial F}{\partial x}(a)$. The advantage of these notations is that if we write $x = (u, v)$, then $F'(a)$ can be decomposed into two block matrices $\begin{bmatrix} F_u(a) & F_v(a) \end{bmatrix} = \begin{bmatrix} \frac{\partial F}{\partial u}(a) & \frac{\partial F}{\partial v}(a) \end{bmatrix}$.

Definition 5.6.5 (Small-Oh Notation). For functions $f : X \to Y$ and $g : X \to Y'$ between normed spaces, the notation $f(h) = o(g(h))$ means for any $\epsilon > 0$, there is $\delta > 0$ such that

$$\|h\| < \delta \implies \|f(h)\| \le \epsilon \|g(h)\|.$$

For example, if a function G is differentiable at a, $\lim_{\|h\|\to 0} \|G(a+h) - G(a) - G'(a)h\|/\|h\| = 0$, so that by definition,

$$G(a+h) - G(a) - G'(a)h = o(h).$$

Proposition 5.6.6 (Chain Rule). *Let $U \subseteq \mathbb{R}^m$ and $V \subseteq \mathbb{R}^k$ be open and consider the composition $U \xrightarrow{G} V \xrightarrow{F} \mathbb{R}^n$. If G is differentiable at $a \in U$ and F is differentiable at $G(a)$, then $F \circ G$ is differentiable at a, moreover,*

$$D(F\circ G)(a) = DF(G(a)) \circ DG(a).$$

Proof. By hypothesis $G(a+h) = G(a) + DG(a)h + o(h)$ and $F(G(a)+k) = F(G(a)) + DF(G(a))k + o(k)$, hence

$$F\circ G(a+h) - F\circ G(a) - DF(G(a))\circ DG(a)(h)$$

[12]We explain it in the second remark following this definition

[13]When we are going to do computation, all vectors are understood to be represented by column vectors.

$$
\begin{aligned}
&= F\Big(G(a)+[DG(a)(h)+o(h)]\Big) - F(G(a)) - DF(G(a)) \circ DG(a)(h) \\
&= DF(G(a))(o(h)).
\end{aligned}
$$

Since $\|DF(G(a))(o(h))\| \le \|DF(G(a))\|\|o(h)\|$, showing that $F \circ G$ is differentiable at a. □

The linearity of "taking derivative" can be similarly proved.

Remark. Let $F : \mathbb{R}^n \to \mathbb{R}^m$, for $x, h \in \mathbb{R}^n$, the vector $F'(x)(h)$ is actually a directional derivative. To see this, for $t \in \mathbb{R}$, let $\Psi(t) = x + th$, then $\Psi(0) = x$ and $\Psi'(0) = h$, so

$$
F'(x)(h) = F'(\Psi(0))\Psi'(0) = (F \circ \Psi(t))'(0) = \lim_{t\to 0} \frac{F(x+th) - F(x)}{t}.
$$

In other words, $F'(x)(h)$ is the derivative of F at x along the direction h. If F does parametrize a surface, then $F'(x)(h)$ will be a tangent vector at x.

Proposition 5.6.7. *If all partial derivatives $\frac{\partial f_i}{\partial x_j}$ of $F : \mathbb{R}^n \to \mathbb{R}^m$ exist near a point a and are continuous at a, then F is differentiable at a.*

Proof. By the first remark following Definition 5.6.3 it suffices to show when the partial derivatives of $f : \mathbb{R}^n \to \mathbb{R}$ exist near a and are continuous at a, then f is differentiable at a. To see this, we observe that

$$
\begin{aligned}
& f(x_1, x_2, \ldots, x_n) - f(a_1, a_2, \ldots, a_n) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) \\
&= [f(x_1, x_2, \ldots, x_n) - f(a_1, x_2, \ldots, x_n)] + [f(a_1, x_2, \ldots, x_n) - f(a_1, a_2, \ldots, x_n)] \\
&\quad + \cdots + [f(a_1, \ldots, a_{n-1}, x_n) - f(a_1, \ldots, a_n)] - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) \\
&= \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}(y^{(i)}) - \frac{\partial f}{\partial x_i}(a) \right)(x_i - a_i),
\end{aligned}
$$

where $y^{(i)} = (a_1, \ldots, a_{i-1}, c_i, x_{i+1}, \ldots, x_n)$, for some c_i between a_i and x_i, hence $\|y^{(i)} - a\|_2 \le \|x - a\|_2$. As $x \to a$, $y^{(i)} \to a$ for each i and since $|x_i - a_i| \le \|x - a\|_2$, the result follows from the continuity of $\frac{\partial f}{\partial x_i}$'s at a. □

Remark. The condition of Proposition 5.6.7 can be slightly weakened. For example, consider $X \subseteq \mathbb{R}^2$ and $f : X \to \mathbb{R}$, if $\frac{\partial f}{\partial x}(x_0, y_0)$ exists and $\frac{\partial f}{\partial y}$ is continuous at (x_0, y_0), then f is differentiable at (x_0, y_0). In general we can replace one of the continuity of $\frac{\partial f}{\partial x_i}$'s at some point by merely differentiability at some point.

It is a warning that although $DF(a)$ exists implies $F'(a)$ exists, the converse can be false. The former one implies a linear transform T exists, denoted by $DF(a)$, such that (5.6.4) holds, but the latter one merely implies all partial derivative exists.

Example 5.6.8. Consider map $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$
f(x,y) = \begin{cases} \dfrac{xy}{x^2 + y^2}, & (x,y) \ne 0, \\ 0, & (x,y) = (0,0). \end{cases}
$$

f' exists at (0,0) and $f'(0,0) = \begin{bmatrix} 0 & 0 \end{bmatrix}$, but f is not continuous at (0,0), hence cannnot be differentiable there. $\square$

Next, the converse of Proposition 5.6.7 can also be false:

Example 5.6.9. Consider $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\dfrac{1}{x^2+y^2}, & (x,y) \neq 0, \\ 0, & (x,y) = (0,0). \end{cases}$$

f is differentiable at (0,0) but $f'(x,y)$ is not continuous there. $\square$

A little summary is given in Figure 5.10.

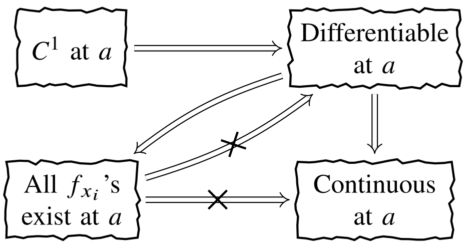


Figure 5.10: A brief relation.

Definition 5.6.10. Let $U \subseteq \mathbb{R}^n$ be open, a function $F : U \to \mathbb{R}^m$ is said to be **continuously differentiable** if F' exists and continuous on U.

Remark. Note that $F' : x \mapsto F'(x)$ is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^{m\times n}$, in order to describe the continuity of this map we need to choose a norm on $\mathbb{R}^{m\times n}$. Let's choose $\|\cdot\|_\infty$, then F' is continuous at a iff $\lim\limits_{x\to a}\|F'(x)-F'(a)\|_\infty = 0$ iff $\lim\limits_{x\to a}\max\limits_{1\le i\le m}\sum_{j=1}^n |\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)| = 0$ iff $\lim\limits_{x\to a}|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)| = 0$ for all i,j iff all partial derivatives are continuous at a.

Hence F is continuously differentiable on an open set U iff all partial derivatives exist on U and are continuous on U, or equivalently, F is differentiable on U and $F'(x)$ is continuous on U, in that case we may abbreviate it as "***F* is C^1 on *U***".

5.6.2 Inverse and Implicit Function Theorem

In the sequel we will be preparing for the proof of inverse function theorem.

Proposition 5.6.11. *Let $F : [a,b] \to \mathbb{R}^n$ be continuous and differentiable on (a,b), then there is $x \in (a,b)$ so that*

$$\|F(b) - F(a)\|_2 \le \|F'(x)\|_2(b-a).$$

Proof. Let $z = F(b) - F(a)$. Define $g(t) = z\cdot F(t)$, then $g : (a,b) \to \mathbb{R}$ is differentiable on (a,b) and hence by mean-value theorem for real-valued functions on an interval, there is $x \in (a,b)$ such that $g(b) - g(a) = g'(x)(b-a) = z \cdot F'(x)(b-a)$. Since $g(b) - g(a) = \|z\|_2^2$, by Cauchy-Schwarz inequality,

$$\|z\|_2^2 = \|z \cdot F'(x)\|_2(b-a) \le \|z\|_2 \|F'(x)\|_2(b-a). \qquad \square$$

Proposition 5.6.11 is an analogy of mean-value theorem of real-valued function on real line, it can be directly extended to differentiable functions $F : \mathbb{R}^m \to \mathbb{R}^n$:

Theorem 5.6.12. *Let $U \subseteq \mathbb{R}^m$ be open and convex and $F : U \to \mathbb{R}^n$ a differentiable function on U. For every $a, b \in U$, there is x on the segment joining a and b such that*

$$\|F(b) - F(a)\|_2 \le \|F'(x)\|_2 \|b-a\|_2.$$

Proof. Let $a, b \in U$, define $G(t) = F((1-t)a + tb)$, then $G : [0,1] \to \mathbb{R}^n$ is differentiable on $(0,1)$, by Proposition 5.6.11 there is t_0 such that

$$\|G(1) - G(0)\|_2 \le \|G'(t_0)\|_2.$$

Since $G(1) = F(b), G(0) = F(a)$ and $G'(t_0) = F'((1-t_0)a + t_0 b)(b-a)$, by letting $x = (1-t_0)a + t_0 b$,

$$\|F(b) - F(a)\|_2 \le \|F'(x)(b-a)\|_2 \le \|F'(x)\|_2 \|b-a\|_2. \qquad \square$$

It is useful to keep Theorem 5.6.12 in mind when constructing the following important class of mappings in analysis.

Definition 5.6.13. Let X, Y be metric spaces, a map $F : X \to Y$ is said to be a **contraction** if there is $c \in [0,1)$ such that for every $x, y \in X$,

$$d(F(x), F(y)) \le c\, d(x,y). \tag{5.6.14}$$

Recall that a fixed point x of F is an element that satisfies $F(x) = x$.

Theorem 5.6.15 (Contraction Mapping). *Let X be a complete metric space and $F : X \to X$ a contraction, then F has a unique fixed point in X.*

Proof. Pick $x_0 \in X$, define inductively that $x_n = F(x_{n-1}) \in X$, for $n = 1, 2, \ldots$. Let c be a constant such that (5.6.14) holds for all $x, y \in X$. Then for each n, $d(x_{n+1}, x_n) = d(F(x_n), F(x_{n-1})) \le c\, d(x_n, x_{n-1})$, hence $d(x_{n+1}, x_n) \le c^n d(x_1, x_0)$. A usual telescoping technique shows that for $m > n$,

$$\begin{aligned} d(x_m, x_n) &\le d(x_m, x_{m-1}) + d(x_{m-1}, x_n) \\ &\le \cdots \\ &\le (c^{m-1} + c^{m-2} + \cdots + c^n) d(x_1, x_0) \\ &\le \frac{c^n}{1-c} d(x_1, x_0), \end{aligned}$$

showing that $\{x_n\}$ is a Cauchy sequence, hence $\{x_n\}$ converges to $x \in X$. But then $x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} F(x_n) = F(x)$, so x is a fixed point of F. The fixed point must be unique since F is a contraction. $\square$

Proposition 5.6.16. *Let $T \in GL(n, \mathbb{R})$, if $\|S\|_2 \le \|T^{-1}\|_2^{-1}$, then $T - S \in GL(n, \mathbb{R})$.*

Proof. Since $T \in GL(n,\mathbb{R})$, $T - S = T(I - T^{-1}S)$ is invertible iff $I - T^{-1}S$ does, a direct computation shows that $\sum_{n\geq 0}(T^{-1}S)^n$ (which converges absolutely by hypothesis) is continuous and inverse to $I - T^{-1}S$. □

Remark. Proposition 5.6.16 proves that $GL(n,\mathbb{R})$ is an open subset of $\mathbb{R}^{n\times n}$. Another form of Proposition 5.6.16 is that given T invertible, if A is a matrix such that $\|A - T\|_2 < \|T^{-1}\|_2^{-1}$, then A is also invertible.

Theorem 5.6.17 (Inverse Function). *Let $O \subseteq \mathbb{R}^n$ be open and $F : O \to \mathbb{R}^n$ continuously differentiable on O. If $F'(a)$ is invertible for some $a \in O$, then:*

(i) There are open sets $U \subseteq O$ and V such that $a \in U$, $F(a) \in V$, F is injective on U and $F(U) = V$, and;

(ii) If $G : V \to U$ is the inverse of $F|_U$ (exists by (i)), then G is continuously differentiable on V.

Proof. (i) Put $A = F'(a)$. Let $y \in \mathbb{R}^n$, define $\Phi_y(x) = x + A^{-1}(y - F(x))$. We note that $y = F(x)$ iff x is a fixed point Φ_y. For each $y \in \mathbb{R}^n$, $\Phi'_y(x) = I - A^{-1}F'(x) = A^{-1}(A - F'(x))$. In view of Theorem 5.6.12, we can choose an open ball U containing a small enough such that

$$x \in U \implies \|A - F'(x)\|_2 < \frac{1}{2}\|A^{-1}\|_2^{-1}, \tag{5.6.18}$$

then $x \in U$ implies $\|\Phi'_y(x)\|_2 \leq \|A^{-1}\|_2\|A - F'(x)\|_2 < \frac{1}{2}\|A^{-1}\|_2\|A^{-1}\|_2^{-1} = \frac{1}{2}$, from that we conclude for every $u,v \in U$ and $y \in \mathbb{R}^n$,

$$\|\Phi_y(u) - \Phi_y(v)\|_2 \leq \frac{1}{2}\|u - v\|_2.$$

Hence Φ_y is a contraction on U, fixed point is unique, namely, there is at most one $x \in U$ such that $F(x) = y$, showing that F is injective on U.

Next we try to show $V := F(U)$ is open in $\mathbb{R}^n$. Let $y_0 = F(x_0)$, for some $x_0 \in U$. Let $B := B(x_0,r)$ be such that $\overline{B} \subseteq U$. Now Φ_y is a contraction on U and for each $x \in \overline{B}$,

$$\begin{aligned}\|\Phi_y(x) - x_0\|_2 &\leq \|\Phi_y(x) - \Phi_y(x_0)\|_2 + \|\Phi_y(x_0) - x_0\|_2\\ &\leq \frac{r}{2} + \|A^{-1}\|_2\|y - y_0\|_2,\end{aligned}$$

hence if we choose y such that $\|y - y_0\|_2 < \frac{r}{2\|A^{-1}\|_2}$, then $\|\Phi_y(x) - x_0\|_2 \leq r$, and hence $\Phi_y : \overline{B} \to \overline{B}$ is a contraction with a fixed point $x \in \overline{B}$ by Theorem 5.6.15, so that $y = F(x) \in V$, thus V is open.

(ii) Let $y, y + k \in V$ and $x, x + h \in U$ such that $F(x) = y$ and $F(x + h) = y + k$. Then by $\|\Phi_y(x + h) - \Phi_y(x)\|_2 = \|h - A^{-1}k\|_2 \leq \frac{1}{2}\|h\|_2$, we have

$$\|h\|_2 \leq 2\|A^{-1}\|_2\|k\|_2. \tag{5.6.19}$$

Let $G : V \to U$ be the inverse of $F|_U$. By (5.6.18) and Proposition 5.6.16, $F'(x)^{-1}$ exists for each $x \in U$, while by the formula $(f^{-1})'(y) = f'(f^{-1}(y))^{-1}$ we learnt in calculus, a natural candidate of linear approximation of G at y is

$$z \mapsto G(y) + F'(G(y))^{-1}(z - y). \tag{5.6.20}$$

We now check that (5.6.20) is a correct choice. Write

$$\|G(y+k)-G(y)-F'(G(y))^{-1}(k)\|_2 = \|F'(x)^{-1}(F(x+h)-F(x)-F'(x)h)\|_2,$$

by using (5.6.19) and differentiability of F at x, we conclude that G is differentiable at each $y \in V$. Finally since $G'(y) = F'(G(y))^{-1}$, F',G are continuous and inverse of a matrix with continuous entries is also continuous, hence $G'(y)$ is continuous. $\square$

A map $F : X \to Y$ is said to be a diffeomorphism if F is bijective and both F and F^{-1} are differentiable. We say that $F : X \to Y$ is a C^1-diffeomorphism if both F and F^{-1} are continuously differentiable. Theorem 5.6.17 says that if a map $F : U \to \mathbb{R}^n$ is continuously differentiable on an open $U \subseteq \mathbb{R}^n$, then

$$F'(x) \text{ invertible} \implies F \text{ is a } C^1\text{-diffeomorphism near } x.$$

Corollary 5.6.21. *Let $U \subseteq \mathbb{R}^n$ be open and $F : U \to \mathbb{R}^n$ injective, continuously differentiable with $\det(F'(x)) \neq 0$ for every $x \in U$, then the following holds:*

(i) $F(U)$ is open in $\mathbb{R}^n$.

(ii) F^{-1} exists and is continuously differentiable

Next we also mention a simple consequence of inverse function theorem which will not be used in this chapter.

Theorem 5.6.22 (Implicit Function). *Let O be open in $\mathbb{R}^{m+n}$ and $F : O \to \mathbb{R}^n$ continuously differentiable. Let $(x,y) \in \mathbb{R}^m \times \mathbb{R}^n$, if there is $(a,b) \in \mathbb{R}^m \times \mathbb{R}^n$ such that $F(a,b) = 0$ and $\det(F_y(a,b)) \neq 0$, then there are an open set U containing a and a unique map $G : U \to \mathbb{R}^n$ such that for every $u \in U$:*

(i) $G(a) = b$ and $F(u,G(u)) = 0$.

(ii) $G'(u) = -[F_y(u,G(u))]^{-1}F_x(u,G(u))$.

Let's denote $\pi : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ the canonical projection map.

Proof. (i) Define $H : O \to \mathbb{R}^{m+n}$ by $(x,y) \mapsto (x,F(x,y))$, then for $(u,v) \in \mathbb{R}^m \times \mathbb{R}^n$,

$$H'(u,v) = \left[\begin{array}{c|c} I & 0 \\ \hline F_x(u,v) & F_y(u,v) \end{array}\right], \tag{5.6.23}$$

hence H is continuously differentiable on O and $\det(H'(a,b)) \neq 0$, by inverse function theorem H is C^1-diffeomorphic on some open V, $(a,b) \in V$. Write $H^{-1} = (S,T) : H(V) \to V$, and note that

$$(x,y) = H^{-1}(H(x,y)) = (S(x,F(x,y)),T(x,(F(x,y))),$$

therefore for convenience we choose $S(x,F(x,y)) = x$ for $(x,y) \in V$.

For $(x,y) \in V$, we note that $F(x,y) = 0$ iff $H(x,y) = (x,0)$ iff $(x,y) = H^{-1}(x,0)$ iff $(x,y) = (x,T(x,0))$ iff $y = T(x,0)$. If we define $G(x) = T(x,0)$, then for each $x \in \pi(V) =: U$, $(x,G(x))$ is the unique solution to $F(x,y) = 0$. As T is continuously differentiable[14], so is G, thus we are done.

[14] i.e., by the remark following Definition 5.6.10, all partial derivatives exist and are continuous.

(ii) Finally by differentiating the equality $F(u,G(u)) = 0$ on U, we have

$$G'(u) = -[F_y(u,G(u))]^{-1} F_x(u,G(u)). \tag{5.6.24}$$

Here F_y is invertible due to (5.6.23) and the fact that $(u,G(u)) \in V$. □

Remark. From the proof, the map (id,G) maps U onto $V \cap F^{-1}(0)$. We may say that (id,G) is a local "parametrization" of $F^{-1}(0)$ at $(a,b) \in F^{-1}(0)$. Note that if F has partial derivatives of any order, so is G by the equation (5.6.24). Moreover, $(\mathrm{id},G)'(x)$ always has full rank.

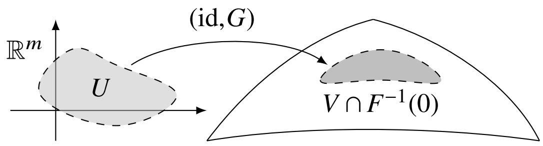


Figure 5.11: A local parametrization.

Remark. For $F : \mathbb{R}^{m+n} \to \mathbb{R}^m$, replace the condition $\det(F_y(a,b)) \neq 0$ by $\det(F_x(a,b)) \neq 0$ in implicit function theorem, we have an analogous statement that there is a unique map $G : V \to \mathbb{R}^m$ on an open V containing b such that for every $u \in V$:

(i) $G(b) = a$ and $F(G(u),u) = 0$.

(ii) $G'(u) = -[F_x(G(u),u)]^{-1} F_y(G(u),u)$.

Indeed, this results from renaming the coordinates. Let v correspond to x and u correspond to y, i.e., $v \in \mathbb{R}^m$ and $u \in \mathbb{R}^n$. Let $A \in \mathbb{R}^{(m+n)\times(m+n)}$ be such that $A(u,v) = (v,u)$. Define $\tilde{F} = F \circ A$, then

$$J\tilde{F}(u,v) = J(F \circ A)(u,v) = JF(A(u,v))JA(u,v) = JF(v,u)A = \begin{bmatrix} F_y(v,u) & F_x(v,u) \end{bmatrix},$$

so that $\tilde{F}_u(u,v) = F_y(v,u)$ and $\tilde{F}_v(u,v) = F_x(v,u)$. Now $\tilde{F}_v(b,a) = F_x(a,b)$ is invertible, so that locally v can be "solved" in terms of u, i.e., we can find open V containing b and a continuously differentiable G on V such that $(u,G(u))$ solves $\tilde{F}(u,v) = 0$ (iff $F(G(u),u) = 0$), $G(b) = a$ and

$$G'(u) = -[\tilde{F}_v(u,G(u))]^{-1} \tilde{F}_u(u,G(u)) = -[F_x(G(u),u)]^{-1} F_y(G(u),u).$$

Example 5.6.25. Let $f(x_1,\ldots,x_k)$ be a homogeneous polynomial of degree d, i.e., $f(tx_1,\ldots tx_k) = t^d f(x_1,\ldots,x_k)$. If $c \neq 0$, then $f^{-1}(c) \subseteq \mathbb{R}^k$ can be locally parametrized by a smooth (i.e., has partial derivatives of any order) function. To see this, let $f(a) = c$, we claim that one of $\frac{\partial f}{\partial x_i}(a)$'s must be nonzero. If $\nabla f(a) = 0$, then consider $F : \mathbb{R} \to \mathbb{R}$ defined by $F(t) = f(ta) = t^d f(a)$, one has

$$dt^{d-1} f(a) = F'(t) = \nabla f(ta) \cdot a,$$

if we take $t = 1$,

$$df(a) = 0 \cdot a = 0,$$

a contradiction, hence $\frac{\partial f}{\partial x_i}(a) \neq 0$ for some i. By implicit function theorem x_i can be "solved" (implicitly) in terms of x_j, $j \neq i$, i.e., there is a continuously differentiable $g : U \to \mathbb{R}$ on some open $U \subseteq \mathbb{R}^{k-1}$ containing $(a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_k)$ such that $(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_k) \in U$ implies

$$f(x_1,\ldots,x_{i-1},g(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_k),x_{i+1},\ldots,x_k) = c.$$

Finally smoothness of the mapping results from the formula of derivatives of g. □

5.7 Integration on $\mathbb{R}^n$

5.7.1 Linear Change of Variable

Consider a real matrix $T \in GL(n,\mathbb{R})$[15], the Borel measure on $\mathbb{R}^n$ defined by

$$\mu_T(E) = m(T(E)) \tag{5.7.1}$$

is also σ-finite and translation invariant (recall that T is a Lipschitz map). It is interesting to see how the measure of E is changed under the linear transform T.

Theorem 5.7.2. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transform and $E \in \mathcal{L}_n$, then $T(E) \in \mathcal{L}_n$ and*

$$m(T(E)) = |\det T| m(E). \tag{5.7.3}$$

By Proposition 5.5.14 a Lipschitz map takes a measurable set to a measurable set and also a set of measure zero to a set of measure zero. To complete the proof of Theorem 5.7.2, it remains to show (5.7.3) holds when E is Borel, and the result can be directly extended by using (ii) of Proposition 5.5.6.

Proof. Let $E \in \mathcal{B}_{\mathbb{R}^n}$. If $\det T = 0$, we leave it as an exercise to show that subspace of $\mathbb{R}^n$ of dimension less than n has m-measure zero.

Assume $|\det T| > 0$. Define μ_T as in (5.7.1), then since μ_T is a Borel measure that is σ-finite and translation invariant, by (iii) of Theorem 5.5.11 there is a constant $C(T) \geq 0$ such that $\mu_T(A) = m(T(A)) = C(T)m(A)$ for each $A \in \mathcal{B}_{\mathbb{R}^n}$. We need to show $C(T) = |\det(T)|$.

Let $H,K \in GL(n,\mathbb{R})$, the following computation

$$C(HK)m(A) = m(H(K(A))) = C(H)m(K(A)) = C(H)C(K)m(A),$$

holds for each $A \in \mathcal{B}_{\mathbb{R}^n}$, hence $C(HK) = C(H)C(K)$. To evaluate $C(T)$, we use **??** which asserts that there are orthogonal matrices $U,V \in \mathbb{R}^{n\times n}$ and a diagonal matrix $\Sigma \in \mathbb{R}^{n\times n}$ such that $T = U\Sigma V^t$[16]. Now $C(U) = C(V^t) = 1$ since orthogonal matrices leave the sphere $\{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$ unchanged. Consider the cube $Q := [0,1]^n$, a direct computation shows that

$$C(\Sigma) = C(\Sigma)m(Q) = m(\Sigma(Q)) = |\sigma_1 \cdots \sigma_n| = |\det \Sigma|,$$

hence $C(T) = C(U)C(\Sigma)C(V^t) = |\det U||\det \Sigma||\det V^t| = |\det T|$. □

[15] It is the collection of real invertible matrices, called general linear group.

[16] It is called singular value decomposition. Conventionally the diagonal values of Σ are denoted by $\sigma_1, \ldots, \sigma_n > 0$, called singular values.

Theorem 5.7.4. *Let $T \in GL(n,\mathbb{R})$, if f is Lebesgue measurable on $\mathbb{R}^n$, so is $f \circ T$. If $f \geq 0$ or $f \in L^1(m)$, then*

$$\int_{\mathbb{R}^n} f \circ T\, dm = \frac{1}{|\det T|} \int_{\mathbb{R}^n} f\, dm. \tag{5.7.5}$$

Proof. In both cases it suffices to show (5.7.5) holds when $f = \chi_A$, where A is Lebesgue measurable, and this follows directly from Theorem 5.7.2. □

Example 5.7.6. We are going to explain the geometrical meaning of determinant. Let $v_1, v_2, \ldots, v_n$ be column vectors in $\mathbb{R}^n$, then the "parallelepiped" spanned by $\{v_i\}$ is

$$\left\{\sum_{i=1}^n x_i v_i : x_1,\ldots,x_n \in [0,1]\right\} = \left[v_1|\cdots|v_n\right]([0,1]^n),$$

hence

$$m\left(\left\{\sum_{i=1}^n x_i v_i : x_i \in [0,1]\right\}\right) = \left|\det\left[v_1|\cdots|v_n\right]\right|$$

by Theorem 5.7.2. That is, absolute value of determinant is the n-dimensional volume of the parallelepiped spanned by each of its column vectors. □

5.7.2 Nonlinear Change of Variable

We now extend Theorem 5.7.4 with T replaced by nice enough nonlinear functions.

Definition 5.7.7. Let U be an open set in $\mathbb{R}^n$. A function $G : U \to \mathbb{R}^n$ is said to be a **change of variable** if it is injective, continuously differentiable on U and $\det(F'(x)) \neq 0$ for every $x \in U$.

By inverse function Theorem 5.6.17, $G : U \to G(U)$ is a change of variable (on U) if and only if G and G^{-1} are both continuously differentiable, i.e., G is a change of variable if and only if G^{-1} does, and $(G^{-1})'(x) = [G'(G^{-1}(x))]^{-1}$.

Theorem 5.7.8. *Let U be an open set in $\mathbb{R}^n$ and $G : U \to \mathbb{R}^n$ a change of variable. If f is a Lebesgue measurable function on $G(U)$, then $f \circ G$ is Lebesgue measurable on U and if $f \geq 0$ or $f \in L^1(G(U),m)$, then*

$$\int_{G(U)} f(x)\, dm(x) = \int_U f \circ G(x) |\det G'(x)|\, dm(x). \tag{5.7.9}$$

In the simplest case $f = \chi_A$, where $A \subseteq G(U)$, say $A = G(E)$ with $E \subseteq U$ Lebesgue measurable, the equality (5.7.9) becomes

$$m(G(E)) = \int_E |\det G'|\, dm. \tag{5.7.10}$$

This suggests to prove the general case, we may try to prove (5.7.10) first when E is a measurable subset of U. To do this, we first prove the case E is a cube, by experience.

Note that G is a Lipschitz function on each cube contained in U (we shall see this in the proof). Hence G must take a measurable subset of U to a measurable subset of $G(U)$ so that it makes sense to write (5.7.10).

Proof. Let's denote $\|\cdot\| = \|\cdot\|_\infty$ and let $T \in GL(n,\mathbb{R})$. $\|T\|$ is the "maximum absolute row sum" that satisfies $\|Tx\| \le \|T\|\|x\|$. Let $Q \subseteq U$ be a cube with center $a \in U$ and $2h$ side length, i.e., $Q = \{x \in \mathbb{R}^n : \|x-a\| \le h\}$.

Let $x \in Q$, write $G = (g_1,\dots,g_n)$. Let i be fixed, by mean-value theorem there is a y lying on the segment joining x and a such that $g_i(x)-g_i(a) = \sum_{j=1}^n \frac{\partial g_i}{\partial x_j}(y)(x_j - a_j)$. Since G' is continuous,

$$|g_i(x)-g_i(a)| \le h\sum_{j=1}^n \left|\frac{\partial g_i}{\partial x_j}(y)\right| \le h\|G'(y)\| \le h\sup_{y\in Q}\|G'(y)\|,$$

this is true for each i, hence $\|G(x)-G(a)\| \le h\sup_{y\in Q}\|G'(y)\|$ and

$$m(G(Q)) \le \left(\sup_{y\in Q}\|G'(y)\|\right)^n m(Q).$$

Replace G by $T^{-1}\circ G$, we have

$$m(G(Q)) = |\det T| m(T^{-1}\circ G(Q)) \le |\det T|\left(\sup_{y\in Q}\|T^{-1}\circ G'(y)\|\right)^n m(Q). \tag{5.7.11}$$

Since $G'(x)$ is continuous, for any $\epsilon > 0$ there is a $\delta > 0$ such that when $u,v \in Q$ and $\|u-v\| < \delta$, $\|(G'(u))^{-1}G'(v)\| < 1+\epsilon$. Divide Q further into cubes having disjoint interiors $Q_1,Q_2,\dots,Q_N$ with centers $x_1,x_2,\dots,x_N$ such that the lengths of their sides are at most δ. We replace Q by Q_i and T by $G'(x_i)$ in (5.7.11), having

$$m(G(Q_i)) \le |\det G'(x_i)|\left(\sup_{y\in Q_i}\|[G'(x_i)]^{-1}\circ G'(y)\|\right)^n m(Q_i) \le (1+\epsilon)^n|\det G'(x_i)| m(Q_i),$$

this implies

$$m(G(Q)) \le \sum_{i=1}^n m(G(Q_i)) \le (1+\epsilon)^n \int_Q \sum_{i=1}^N |\det G'(x_i)|\chi_{Q_i}(x)\, dm(x).$$

For each $\varepsilon > 0$ we can find even smaller δ at the beginning such that $|\det G'(x) - \det G'(y)| < \varepsilon$ whenever $x,y \in Q$ and $\|x-y\| < \delta$. For this fixed δ, for a.e. $x \in Q$, $x \in Q_i$ for some unique i and hence $\sum_{i=1}^N |\det G'(x_i)|\chi_{Q_i}(x) = |\det G'(x_i)| < |\det G'(x)| + \varepsilon$, hence $m(G(Q)) \le (1+\epsilon)^n \int_Q (|\det G'(x)|+\varepsilon)\, dm(x)$. Since this is true for every $\varepsilon > 0$ and also every $\epsilon > 0$, we conclude

$$m(G(Q)) \le \int_Q |\det G'(x)|\, dm(x).$$

Since U is a countable union of cubes $\{R_i\}$, with disjoint interiors, write $U = \bigcup R_i$, and then by monotone convergence theorem,

$$m(G(U)) \le \sum_{i=1}^\infty m(G(R_i)) \le \sum_{i=1}^\infty \int_{R_i} |\det G'|\, dm = \int_U |\det G'|\, dm.$$

Also if E is any bounded Lebesgue measurable subset of U, there is a descending collection of bounded open sets $O_i \supseteq E$ such that $O_i \subseteq U$ and $m(\bigcap O_i - E) = 0$. Hence by continuity of integration,

$$m(G(E)) \le m\left(\bigcap G(O_i)\right) = \lim m(G(O_i)) \le \lim \int_{O_i} |\det G'|\, dm = \int_E |\det G'|\, dm. \tag{5.7.12}$$

If E is unbounded, we replace E in (5.7.12) by $E \cap [-N,N]^n$ and use continuity of measure and monotone convergence theorem to conclude (5.7.12) in general. At this point it is easy to show (5.7.9) is true if "=" is replaced by "≤" when f is nonnegative simple function. Hence for all $f \geq 0$,

$$\int_{G(U)} f(x)\, dm(x) \leq \int_U f \circ G(x) |\det G'(x)|\, dm(x).$$

Replace f by a function $g \geq 0$ on U, replace the role of G by G^{-1} and replace the role of $G(U)$ by U, then by the same reasoning we have

$$\int_U g(x)\, dm(x) \leq \int_{G(U)} g \circ G^{-1}(x) |\det(G^{-1})'(x)|\, dm(x),$$

we substitute $g = f \circ G \cdot |\det G'|$ to get

$$\int_U f \circ G(x) |\det G'(x)|\, dm(x) \leq \int_{G(U)} f(x)\, dm(x).$$

Hence the case that $f \geq 0$ is done. In the case that f is integrable, we split $\operatorname{Re} f$ and $\operatorname{Im} f$ into positive and negative parts. □

Example 5.7.13 (Polar Coordinate in $\mathbb{R}^2$). Let's denote $m_2 = A$, called area measure. We let $x = r\cos\theta$, $y = r\sin\theta$, i.e., $(x,y) = G(r,\theta)$, where

$$G : (0,\infty) \times (0,2\pi) \to \mathbb{R}^2 - [0,\infty) \times \{0\}; \quad (r,\theta) \mapsto (r\cos\theta, r\sin\theta).$$

G is injective, C^1 and $|\det G'(r,\theta)| = r \neq 0$. For $f \geq 0$ or $f \in L^1(\mathbb{R}^2, A)$, since $[0,\infty) \times \{0\}$ has A-measure zero, we have

$$\int_{\mathbb{R}^2} f\, dA = \int\limits_{\mathbb{R}^2 - [0,\infty)\times\{0\}} f\, dA = \int\limits_{(0,\infty)\times(0,2\pi)} f(r\cos\theta, r\sin\theta) r\, dA(r,\theta).$$

Replace f by $\chi_E f$, where $E \subseteq \mathbb{R}^2$ is Lebesgue measurable, then

$$\begin{aligned} \int_E f\, dA &= \int\limits_{(0,\infty)\times(0,2\pi)} \chi_{G^{-1}(E)}(r,\theta) f(r\cos\theta, r\sin\theta) r\, dA(r,\theta) \\ &= \int_0^{2\pi} \int_0^{\infty} \chi_{G^{-1}(E)}(r,\theta) f(r\cos\theta, r\sin\theta) r\, dm(r) dm(\theta) \end{aligned} \tag{5.7.14}$$

where $G^{-1}(E) := \{(r,\theta) \in (0,\infty) \times (0,2\pi) : G(r,\theta) \in E\}$[(17)] and finding it becomes a major task. □

Example 5.7.15. As an application we try to compute $\int_{-\infty}^{\infty} e^{-x^2}\, dx$ for which we usually pretended we can do change of variable without any justification in calculus course.

By Fubini-Tonelli theorem as $e^{-x^2-y^2}$ is nonnegative $\mathcal{B}_{\mathbb{R}^2}$-measurable,

$$\left(\int_{\mathbb{R}} e^{-x^2}\, dm(x) \right)^2 = \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-x^2-y^2}\, dm(x) dm(y) = \int_{\mathbb{R}^2} e^{-(x^2+y^2)}\, dA.$$

[(17)] It may be confusing, we emphasize it is not the image of the function G^{-1} on $\mathbb{R}^2$ excluding positive axis.

Let $E = \mathbb{R}^2$ in (5.7.14), $G^{-1}(\mathbb{R}^2) = (0,\infty)\times(0,2\pi)$, yielding

$$\left(\int_{\mathbb{R}} e^{-x^2}\, dm(x)\right)^2 = \int_0^{2\pi}\int_0^{\infty} e^{-r^2} r\, dm(r) dm(\theta).$$

Recall that if f is absolutely Riemann integrable on $\mathbb{R}$, then it is also Lebesgue integrable on $\mathbb{R}$ and the two integrals agree. Hence under Riemann integration,

$$\int_{-\infty}^{\infty} e^{-x^2}\, dx = \sqrt{\int_0^{2\pi}\int_0^{\infty} e^{-r^2} r\, dr\, d\theta} = \sqrt{\pi}. \qquad \square$$

Example 5.7.16. In 5.5.13 we have shown that

$$m_n(B(0,r)) = r^n m_n(B(0,1)).$$

Let's compute the value of $m_n(B(0,1))$. Let G be defined as in Example 5.7.13. In $\mathbb{R}^n$, we define $B_n(r) := B(0,r) := \{x \in \mathbb{R}^n : \|x\|_2 < r\}$ and $V_n(r) = m_n(B_n(r))$. It's known that $V_1(1) = 2, V_2(1) = \pi$. Consider $n \geq 3$, we make use of the complete version of Fubini-Tonelli theorem. In general $\overline{m_h \times m_k} = m_{h+k}$ (why?), we have

$$\begin{aligned}
V_n(1) &= \int_{\mathbb{R}^n} \chi_{B_n(1)}(x)\, dm_n(x) = \int_{\mathbb{R}^2}\int_{\mathbb{R}^{n-2}} \chi_{B_n(1)}(x,u,v)\, dm_{n-2}(x) dm_2(u,v) \\
&= \int_{\mathbb{R}^2}\int_{\mathbb{R}^{n-2}} \chi_{B_{n-2}(\sqrt{1-u^2-v^2})}(x)\, dm_{n-2}(x) dm_2(u,v) \\
&= \int_{B_2(1)}\int_{\mathbb{R}^{n-2}} \chi_{B_{n-2}(\sqrt{1-u^2-v^2})}(x)\, dm_{n-2}(x) dm_2(u,v) \\
&= \int_{B_2(1)} (1-u^2-v^2)^{(n-2)/2} V_{n-2}(1)\, dA(u,v),
\end{aligned}$$

by (5.7.14) we have

$$V_n(1) = V_{n-2}(1)\int_0^{2\pi}\int_0^{\infty} \chi_{G^{-1}(B_2(1))}(r,\theta)(1-r^2)^{(n-2)/2} r\, dm_2(r,\theta).$$

It is straightforward to see $G^{-1}(B_2(1)) = (0,1)\times(0,2\pi)$, hence

$$V_n(1) = V_{n-2}\int_0^{2\pi}\int_0^{1} (1-r^2)^{(n-2)/2} r\, dr\, d\theta = V_{n-2}(1)\frac{2\pi}{n}.$$

Splitting into two cases,

$$V_n(1) = \begin{cases} \dfrac{\pi^{\frac{n}{2}}}{(\frac{n}{2})!}, & \text{if } n \text{ is even,} \\ \dfrac{2^{n+1}\pi^{\frac{n-1}{2}}(\frac{n+1}{2})!}{(n+1)!}, & \text{if } n \text{ is odd.} \end{cases} \qquad \square$$

Remark. Let $u(x) \leq v(x)$ on $[0,1]$, where $u,v : [0,1] \to [0,\infty)$ are measurable. To evaluate

$$\int_D f\, dA, \quad D = \{(x,y) : y \in [u(x),v(x)], x \in [0,1]\},$$

we can do integrated integral in the following way:

$$\int_{\mathbb{R}^2} \chi_D f \, dA = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} (\chi_D)_x f_x \, dy \right) dx = \int_0^1 \int_{\mathbb{R}} \chi_{D_x} f_x \, dy dx.$$

Now $D_x = [u(x), v(x)]$, hence $\int_D f \, dA = \int_0^1 \int_{u(x)}^{v(x)} f \, dy dx$. We see that the main task is to figure out what is D_x (or D^y) in general.

Remark. For $f \in L^1$, people are used to writing $\int f \, dm_1$ as $\int f \, dx$ without any confusion for the following reasons: (i) By Theorem 5.2.30 all absolutely Riemann integrable functions, say f, are Lebesgue integrable and $\int f \, dm_1(x) = \int f \, dx$; (ii) Writing $dm(r)dm(\theta)dm(\cdots$ is cumbersome.

5.8 Exercises and Problems

Exercises

5.1. Give a complete proof of simple approximation theorem.

5.2. Show that in Egoroff's theorem, the hypothesis "$\mu(X) < \infty$" can be replaced by "$|f_1|, |f_2|, \cdots \le g$, where $g \in L^1(X, \mu)$".

5.3. Show that the statement "If there is a sequence of measurable functions $\{f_n\}$ such that $f_n \to f$ pointwise a.e., then f is measurable." can be false if X is incomplete.

Definition 5.8.1. For a sequence of sets $\{E_n\}$, we define

$$\overline{\lim_{n\to\infty}} E_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k \quad \text{and} \quad \underline{\lim_{n\to\infty}} E_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E_k.$$

5.4. Let $\{E_n\}$ be a sequence of measurable sets in (X, Σ, μ), show that

$$\overline{\lim}_{n\to\infty} \chi_{E_n} = \chi_{\overline{\lim}_{n\to\infty} E_n} \quad \text{and} \quad \underline{\lim}_{n\to\infty} \chi_{E_n} = \chi_{\underline{\lim}_{n\to\infty} E_n}.$$

5.5. Let (X, Σ) be a measurable space, $u, v : X \to \mathbb{R}$ measurable functions on X and $\Phi : \mathbb{R}^2 \to \mathbb{R}$ a continuous function. Define

$$h(x) = \Phi(u(x), v(x))$$

for each $x \in X$, show that $h : X \to \mathbb{R}$ is measurable.

5.6. Let $f : [0,1] \to \mathbb{R}$ be an integrable function. Suppose for every interval $J \subseteq [0,1]$ we have $0 \le \displaystyle\int_J f \, dm \le m(J)$, prove that $0 \le f \le 1$ almost everywhere without using Theorem 5.2.43 nor trying to reprove its statement.

5.7. Let $f, g > 0$ be Lebesgue integrable functions on E such that $fg \ge 1$ and $m(E) = 1$. Prove that

$$\int_E f \, dm \int_E g \, dm \ge 1.$$

5.8. Let (X,Σ,μ) be measure space. Suppose $f_n : X \to [0,\infty]$ is measurable for $n = 1,2,\dots$ and $f_1 \geq f_2 \geq \cdots \geq 0$ and $f_n \to f$ pointwise on X. If $f_1 \in \mathcal{L}^1(X)$, prove that

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

5.9. Let E be a measurable subset of $\mathbb{R}$ and $f \in L^1(E,m)$, define $E_k = \{x \in E : |f(x)| < \frac{1}{k}\}$, show that

$$\lim_{k\to\infty} \int_{E_k} |f| \, dm = 0.$$

5.10. (Generalize Riemann-Lebesgue Lemma Slightly) If $g_n : [a,b] \to \mathbb{R}$ is a sequence of measurable functions such that

(i) $|g_n| \leq M$ on $[a,b]$, for $n = 1,2,\dots$.

(ii) For any $c \in [a,b]$, one has $\displaystyle\lim_{n\to\infty} \int_{[a,c]} g_n \, dm = 0$.

Show that for any $f \in L^1([a,b],m)$,

$$\lim_{n\to\infty} \int_{[a,b]} f g_n \, dm = 0.$$

[**Hint:** Use 5.28.]

5.11. In this exercise we will extend the result in Corollary 5.2.23 to complete the proof of a well-known result that: $f : [a,b] \to \mathbb{R}$ is Riemann integrable iff f is bounded and continuous a.e..

Let $f : [a,b] \to \mathbb{R}$ be bounded and continuous m-a.e.[(18)] on $[a.b]$, here m denotes Lebesgue measure on $\mathbb{R}$.

(a) Let $\{P_n\}_{n\geq 1}$ be any sequence of partitions of $[a,b]$ such that each P_{n+1} refines P_n and $\|P_n\| \to 0$. Let φ_n and ψ_n ($\varphi_n \leq f \leq \psi_n$) be defined as in Theorem 5.2.21. Let $x \in (a,b)$ be a point of continuity of f, show that

$$\lim_{n\to\infty} \varphi_n(x) = f(x) = \lim_{n\to\infty} \psi_n(x).$$

(b) Using (a) and the dominated convergence theorem, deduce that

$$\int_{[a,b]} f \, dm = \lim_{n\to\infty} \int_{[a,b]} \varphi_n \, dm = \lim_{n\to\infty} \int_{[a,b]} \psi_n \, dm.$$

(c) Show that f is Riemann integrable on $[a,b]$ and

$$\int_{[a,b]} f \, dm = \int_a^b f(x) \, dx.$$

5.12. Let (X,Σ,μ) be a measure space and $f : X \to [-\infty,\infty]$ integrable over X. Show that for any $\epsilon > 0$, there is a $\delta > 0$ such that for every $A \in \Sigma$,

$$\mu(A) < \delta \implies \left| \int_A f \, d\mu \right| < \epsilon.$$

[(18)] Some property P holds μ-a.e. means P holds except a set of μ-measure zero.

5.13. The purpose of this exercise is to demonstrate that Tonelli's theorem can fail if the σ-finite hypothesis is removed, and also the product measure on $\mathcal{M}\otimes\mathcal{N}$ that extends $A\times B \mapsto \mu(A)\nu(B)$ needs not be unique.

Consider measure space

$$X = ([0,1], \Sigma_X := \{A \in \mathcal{L} : A \subseteq [0,1]\}, m),$$

where $\mathcal{L}$ is Lebesgue σ-algebra and m is Lebesgue measure. Also consider

$$Y = ([0,1], \Sigma_Y := 2^{[0,1]}, c),$$

where c is counting measure. Let $E = \{(x,x) : x \in [0,1]\}$ be the diagonal.

(i) Show that χ_E is $\Sigma_X \otimes \Sigma_Y$-measurable.

(ii) Show that $\int_X \int_Y \chi_E(x,y)\, dc(y) dm(x) = 1$.

(iii) Show that $\int_Y \int_X \chi_E(x,y)\, dm(x) dc(y) = 0$.

(iv) Show that there is more than one measure μ on $\Sigma_X \otimes \Sigma_Y$ with the property that $\mu(E\times F) = m(E)c(F)$ for all measurable rectangles $E\times F \in \Sigma_X \times \Sigma_Y$.

[**Hint:** Use two different ways to perform a double integral to create two different measures.]

5.14. The purpose of this exercise is to demonstrate that Fubini-Tonelli theorem can fail if f is neither nonnegative nor integrable (i.e., one of them is necessary).

Let $X = Y = \mathbb{N}$, $\mathcal{M} = \mathcal{N} = 2^{\mathbb{N}}$ and $\mu = \nu = c$ (counting measure). Define

$$f(m,n) = \begin{cases} 1, & \text{if } m = n, \\ -1, & \text{if } m = n+1, \\ 0, & \text{otherwise.} \end{cases}$$

Show that $\int_{\mathbb{N}\times\mathbb{N}} |f|\, d(\mu\times\nu) = \infty$, while $\int_{\mathbb{N}}\int_{\mathbb{N}} f\, d\mu d\nu$ and $\int_{\mathbb{N}}\int_{\mathbb{N}} f\, d\nu d\mu$ exist and are unequal.

5.15. Let $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ be arbitrary measure spaces (not necessarily σ-finite).

(i) Let $f : X \to \mathbb{C}$ be $\mathcal{M}$-measurable, $g : Y \to \mathbb{C}$ be $\mathcal{N}$-measruable and define $h(x,y) = f(x)g(y)$, show that h is $\mathcal{M}\otimes\mathcal{N}$-measurable.

(ii) If $f \in L^1(\mu)$ and $g \in L^1(\nu)$, then $h \in L^1(\mu\times\nu)$ and

$$\int_{X\times Y} h\, d(\mu\times\nu) = \left(\int_X f\, d\mu\right)\left(\int_Y g\, d\nu\right).$$

5.16. **(Chebychev's Inequality)** Let f be a nonnegative measurable function on X and $\lambda > 0$, show that

$$\mu\{x \in X : f(x) \geq \lambda\} \leq \frac{1}{\lambda}\int_X f\, d\mu.$$

5.17. **(Hölder's Inequality)** Let $a,b \geq 0$ and $p,q \geq 1$ be such that $\frac{1}{p}+\frac{1}{q}=1$. Show that

$$\frac{1}{p}a^p + \frac{1}{q}b^q \geq ab.$$

Hence, or otherwise, show that if $f \in L^p(X,\mu)$, $g \in L^q(X,\mu)$, then $fg \in L^1(X,\mu)$ and

$$\|fg\|_1 \leq \|f\|_p \|f\|_q,$$

and then deduce that for all $a_i, b_i \in \mathbb{C}$,

$$\sum_{k=1}^n |a_k b_k| \leq \sqrt[p]{\sum_{k=1}^n |a_k|^p}\ \sqrt[q]{\sum_{k=1}^n |b_k|^q}.$$

5.18. Let (X,Σ,μ) be a measure space and f an extended real-valued measurable function on X, show that

$$\int_X |f|\,d\mu = 0 \implies f = 0 \text{ a.e. on } X$$

and

$$\int_X |f|\,d\mu < \infty \implies f(x) < \infty \text{ for a.e. } x \in X.$$

5.19. If (X,Σ,μ) is a measure space and if f is μ-integrable, show that for every $\epsilon > 0$ there is $E \in \Sigma$ such that $\mu(E) < \infty$ and $\displaystyle\int_{X-E} |f|\,d\mu < \epsilon$.
[**Hint:** Partition the range of f.]

5.20. Let (X,Σ,μ) be a finite measure space and $1 \leq p \leq q \leq \infty$, show that $L^q(X,\mu) \subseteq L^p(X,\mu)$.

5.21. Suppose that $\{a_{mn}\}_{m,n=1}^{\infty}$ is a double sequence of complex numbers for which at least one of

$$\sum_m \sum_n |a_{mn}| \quad \text{and} \quad \sum_n \sum_m |a_{mn}|$$

is finite, then both of $\sum_m \sum_n a_{mn}$ and $\sum_n \sum_m a_{mn}$ are finite and equal.

5.22. Let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ be σ-algebras on spaces X,Y and Z respectively, show that

$$(\mathcal{A}\otimes\mathcal{B})\otimes\mathcal{C} = \mathcal{A}\otimes(\mathcal{B}\otimes\mathcal{C}) = \mathcal{A}\otimes\mathcal{B}\otimes\mathcal{C}.$$

5.23. Let $f \in L^1(\mathbb{R},m)$, evaluate $\displaystyle\lim_{n\to+\infty}\int_{-\infty}^{\infty} f(x-n)\left(\frac{x}{1+|x|}\right)dm.$

Problems

5.24. Let $f(x) : [a,b] \to (0,\infty)$ and $0 < q \leq b-a$, denote $\Gamma = \{E \subseteq [a,b] : m(E) \geq q\}$, show that

$$\inf_{E\in\Gamma}\left\{\int_E f\,dm\right\} > 0.$$

5.25. Let f be a positive Lebesgue integrable function on $[a,b]$, $\{E_n\}$ a collection of Lebesgue measurable subsets of $[a,b]$. Show that

$$\lim_{n\to\infty}\int_{E_n} f\,dm = 0 \implies \lim_{n\to\infty} m(E_n) = 0.$$

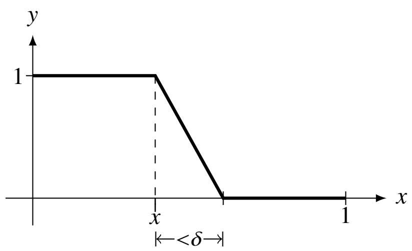
5.26. Let $F, f_1, f_2, \dots \in L^1([0,1], m)$ such that

(i) $|f_n(x)| \le F(x)$ for $n = 1, 2, \dots$.

(ii) $\lim_{n\to\infty}\int_{[0,1]} f_n g\,dm = 0$ for each $g \in C[0,1]$.

Show that for every measurable $E \subseteq [0,1]$, we have $\lim_{n\to\infty}\int_E f_n\,dm = 0$.

[**Hint:** You may use Problem 5.12 and also consider continuous functions of the form:



Conclude your result to finite union of intervals and extend it to open sets, and then extend it to measurable sets.]

5.27. Let $f : [0,1] \to \mathbb{R}$ be a bounded measurable function. Show that

$$\int_{[0,1]} x^n f(x)\,dm = 0 \text{ for } n = 1, 2, \dots \implies f = 0 \text{ a.e. on } [0,1].$$

5.28. Consider the measure space $(\mathbb{R}, \mathcal{L}, m)$. Let $f : \mathbb{R} \to [-\infty, \infty]$ be integrable over $\mathbb{R}$ and $\epsilon > 0$. Establish the following three approximation properties.

(a) There is a simple function η on $\mathbb{R}$ having finite support and $\int_{\mathbb{R}} |f - \eta|\,dm < \epsilon$.

(b) There is a step function s on $\mathbb{R}$ which vanishes outside a closed, bounded interval and $\int_{\mathbb{R}} |f - s|\,dm < \epsilon$.

(c) There is a continuous function g on $\mathbb{R}$ which vanishes outside a bounded set and $\int_{\mathbb{R}} |f - g|\,dm < \epsilon$.

Remark. Now the result can be extended to integration over any measurable subset of $\mathbb{R}$.

5.29. Consider the measure space $(\mathbb{R}, \mathcal{L}, m)$. Let $f : \mathbb{R} \to [-\infty, \infty]$ be integrable over $\mathbb{R}$.

(a) Show that for each t,

$$\int_{-\infty}^{\infty} f(x)\,dm(x) = \int_{-\infty}^{\infty} f(x+t)\,dm(x).$$

(b) Let g be a bounded measurable function on $\mathbb{R}$. Show that

$$\lim_{t\to 0}\int_{-\infty}^{\infty} g(x)\cdot\big(f(x)-f(x+t)\big)\,dm(x)=0.$$

[**Hint:** Density of continuous functions.]

5.30. **(Generalize Riemann-Lebesgue Lemma)** Consider the measure space $(\mathbb{R},\mathcal{L},m)$. Let f be extended real-valued integrable function on $\mathbb{R}$ and g a real-valued bounded integrable function with period $T>0$, then

$$\lim_{t\to+\infty}\int_{\mathbb{R}} f(x)g(tx)\,dm(x)=\frac{1}{T}\int_{[0,T)} g\,dm\int_{\mathbb{R}} f\,dm.$$

[**Hint:** You may need the simple function technique as in Theorem 5.2.15.]

5.31. **(General Lebesgue Dominated Convergence Theorem)** Let (X,Σ,μ) be a measure space. Let $\{f_n\},\{g_n\}$ be sequences of measurable functions on X such that $|f_n|\le g_n$ for each n. Let f,g be measurable functions such that both $f_n\to f, g_n\to g$ pointwise a.e. on X. Show that

$$\lim_{n\to\infty}\int_X g_n\,d\mu=\int_X g\,d\mu<\infty \implies \lim_{n\to\infty}\int_X f_n\,d\mu=\int_X f\,d\mu.$$

5.32. Let $f,f_1,f_2,\dots\in L^p[0,1]$ and $f_n\to f$ a.e., prove that

$$\|f_n-f\|_p\to 0 \iff \|f_n\|_p\to\|f\|_p.$$

5.33. Finish the proof of Proposition 5.2.38.

5.34. Consider (X,Σ,μ) with $\mu(X)<\infty$. Let $f\in L^\infty(X,\mu)$ with $\|f\|_\infty\neq 0$. Prove that

$$\lim_{n\to+\infty}\left(\int_X |f|^n\,d\mu\right)^{1/n}=\lim_{n\to+\infty}\frac{\int_X |f|^{n+1}\,d\mu}{\int_X |f|^n\,d\mu}=\|f\|_\infty.$$

[**Hint:**

(i) Let f be a measurable function on X. A constant M is said to be an essential upper bound of f if $|f|\le M$ a.e.. We define the essential supremum of f by

$$\|f\|_\infty=\inf\{M:|f|\le M \text{ a.e.}\},$$

a simple checking shows that $\|f\|_\infty$ is also an essential upper bound. Hence for any given $\epsilon>0$, $\|f\|_\infty-\epsilon$ cannot be an essential upper bound, i.e., $\mu\{x\in X:|f(x)|>\|f\|_\infty-\epsilon\}>0$.

(ii) One of the limits directly follows from the definition of $\|f\|_\infty$. For another one, you may use Hölder's inequality proved in Problem 5.17.

]

5.35. Let (X,μ) be a measure space and $f:X\to\mathbb{R}$ measurable.

(i) If $\mu(X)<\infty$, show that $f\in L^1(X,\mu)$ if and only if

$$\sum_{n=1}^{\infty}2^n\,\mu(x\in X:|f(x)|\ge 2^n\}<\infty.$$

(ii) If $\mu(X) = \infty$ but f is bounded, show that $f \in L^1(X,\mu)$ if and only if

$$\sum_{n=1}^{\infty} 2^{-n} \mu\{x \in X : |f(x)| \geq 2^{-n}\} < \infty.$$

5.36. Let f be a bounded measurable function on a measure space (X,Σ,μ). Assume that there are constants $C > 0$ and $\alpha \in (0,1)$ such that

$$\mu\{x \in X : |f(x)| > \epsilon\} < \frac{C}{\epsilon^{\alpha}}$$

for every $\epsilon > 0$. Show that $f \in L^1(X,\mu)$.

5.37. Let (X,Σ,μ) be a finite measure space and $f : X \to \mathbb{R}$ a measurable function such that $f^n \in L^1(X,\mu)$ for each $n \in \mathbb{N}$.

(i) If $\lim_{n\to\infty} \int_X f^n \, d\mu$ exists for each n, show that $|f(x)| \leq 1$ for a.e. x.

(ii) Show that there is $c \in \mathbb{R}$ such that $\int_X f^n \, d\mu = c$ for each $n \in \mathbb{N}$ if and only if $f = \chi_A$ a.e. for some $A \in \Sigma$.

5.38. Consider measure spaces (X,Σ,μ) and $(\mathbb{R},\mathcal{B},m)$, where X is σ-finite, $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$ and m is Lebesgue measure. Let $X \times \mathbb{R}$ have the product σ-algebra and $f : X \to [0,\infty)$ be measurable.

(i) Prove that

$$G(f) := \{(x,y) : y \in [0,f(x)], x \in X\} = \bigcup_{x \in X} \{x\} \times [0,f(x)]$$

is measurable, moreover, $\mu \times m(G(f)) = \displaystyle\int_X f \, d\mu$.

(ii) Hence, show that the graph of f, $\Gamma(f) := \{(x,f(x)) : x \in X\}$, has $\mu \times m$-measure zero.

5.39. Prove Proposition 5.4.6.
[**Hint:** Use Corollary 4.3.16.]

5.40. For each $n \in \mathbb{N}$, show that there is a subset of $\mathbb{R}^n$ that is Lebesgue measurable but not Borel.

5.41. (Density of Continuous Functions) By using part (iii) of Proposition 5.5.6, show that for every $f \in L^p(\mathbb{R}^n,m)$ and for every $\epsilon > 0$, there is a continuous function $g : \mathbb{R}^n \to \mathbb{C}$ which vanishes outside a compact set such that

$$\|f-g\|_p := \sqrt[p]{\int_{\mathbb{R}^n} |f-g|^p \, dm} < \epsilon.$$

5.42. Prove Proposition 5.5.14.

5.43. Consider $\|\cdot\| = \|\cdot\|_2$. We note that a map is Lipschitz if and only if its coordinate maps are Lipschitz. For $X \subseteq \mathbb{R}^n$, we try to extend a function $f : X \to \mathbb{R}$ satisfying $\|f(x) - f(y)\| \le L\|x - y\|$ for all $x, y \in X$. Let $a \in X$, show that

$$f_a(x) := f(a) + L\|x - a\|$$

is a Lipscthiz function on $\mathbb{R}^n$. Also show that

$$F(x) := \inf_{a \in X} f_a(x)$$

is a Lipschitz function on the whole $\mathbb{R}^n$ that extends f.

5.44. Show that every subspace of $\mathbb{R}^n$ having dimension less than n is of Lebesgue measure zero.

5.45. Evaluate the integral $\displaystyle\int_0^1 \frac{\tan^{-1} x}{x\sqrt{1-x^2}}\, dx$, where $\tan^{-1}$ is the inverse function of $\tan : (-\pi/2, \pi/2) \to (-\infty, \infty)$.

5.46. Evaluate the integral $\displaystyle\int_0^1 \frac{x^b - x^a}{\ln x}\, dx$, where $a \in (0, b)$, also try to evaluate $\displaystyle\int_0^1 \sin\left(\ln\frac{1}{x}\right) \cdot \frac{x^b - x^a}{\ln x}\, dx$.

5.47. Let $a_i > 0$ for $i = 1, 2, \ldots, n$ and let $J = (0,1) \times \cdots (0,1)$, show that

$$\int_J \frac{1}{x_1^{a_1} + x_2^{a_2} + \cdots + x_n^{a_n}}\, dm(x) < \infty \iff \sum_{i=1}^n \frac{1}{a_i} > 1.$$

[**Hint:** Let $G_k = \{x \in J : x_i^{a_i} \le x_k^{a_k} \text{ for all } i\}$, then $J = \bigcup_{k=1}^n G_k$, show that

$$\int_{G_k} \frac{1}{x_k^{a_k}}\, dm(x) = \int_0^1 x_k^{a_k(\sum_{j=1}^n 1/a_j - 1) - 1}\, dm(x_k)$$

and use the fact that $\displaystyle\int_0^1 t^{s-1}\, dt < \infty$ iff $s > 0$.]

Chapter 6

Signed and Complex Measures, Lebesgue Differentiation Theorem

The aim of this chapter is to discuss countably additive set functions which are not necessarily nonnegative or even real-valued. These arise very naturally and we have encountered some of them before. For example, let $f : (X,\Sigma,\mu) \to [-\infty,\infty]$ or $f : (X,\Sigma,\mu) \to \mathbb{C}$ be integrable, the set function ν defined by

$$\nu(A) = \int_A f \, d\mu \tag{6.0.1}$$

is countably additive and enjoys the continuity of integration property.

Since we will be considering a larger classes of additive set functions, in this chapter, for emphasis, "measures" that are defined in the previous chapters are referred to **positive measures**.

6.1 Signed Measures

6.1.1 Hahn Decompositions

Definition 6.1.1. Let (X,Σ) be a measurable space, a set function $\lambda : \Sigma \to [-\infty,\infty]$ is called a **signed measure** if it has the following properties:

(i) $\lambda(\emptyset) = 0$.

(ii) Either $\lambda(E) < \infty$ for each $E \in \Sigma$ or $\lambda(E) > -\infty$ for each $E \in \Sigma$,

(iii) If $\{E_n\}_{n=1}^{\infty}$ is a disjoint collection of members in Σ, then

$$\lambda\left(\bigsqcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \lambda(E_n). \tag{6.1.2}$$

Since finite signed measures only takes value in $\mathbb{R}$, they are sometimes called **real measures**.

Some remarks are in order. Firstly, property (ii) of Definition 6.1.1 means that if $\nu(A) = \infty$ for some $A \in \Sigma$, then $\nu(E) \neq -\infty$ for each $E \in \Sigma$. Similarly, if $\nu(A) = -\infty$ for some $A \in \Sigma$, then $\nu(E) \neq \infty$ for each $E \in \Sigma$.

Secondly, property (iii) of Definition 6.1.1 means that if $|\lambda(\bigsqcup_{n=1}^{\infty} E_n)| < \infty$, then due to the set equality $\bigsqcup_{n=1}^{\infty} E_n = \bigsqcup_{n=1}^{\infty} E_{\sigma(n)}$ for any bijection $\sigma : \mathbb{N} \to \mathbb{N}$, the value of the series must be independent of any rearrangement. In other words, in case $|\lambda(\bigsqcup_{n=1}^{\infty} E_n)| < \infty$, the convergence in RHS of (6.1.2) *must be absolute*.

Thirdly, by measures we refer to positive or signed measures. Positive measures is a proper subclass of signed measures.

Definition 6.1.3. A signed measure λ on (X,Σ) is said to be **finite** if $|\lambda(X)| < \infty$, and **σ-finite** if there are $X_n \in \Sigma$ such that $X = \bigcup X_n$ and $|\lambda(X_n)| < \infty$.

Example 6.1.4. If μ_1 and μ_2 are positive measures on Σ and at least one of them is finite, then $\nu := \mu_1 - \mu_2$ is a signed measure. □

Although Example 6.1.4 is simple, it is significant because signed measures will be proved to be a difference of two positive measures, with one of them being finite. A precise statement of this result will be made in Jordan decomposition theorem.

Proposition 6.1.5. *Let λ be a signed measure on (X,Σ).*

(i) If $A \in \Sigma$ and $|\lambda(A)| < \infty$, then for any measurable $B \subseteq A$, $|\lambda(B)| < \infty$.

(ii) λ is finite iff $|\lambda(E)| < \infty$ for each $E \in \Sigma$.

Proof. (i) Let $B \subseteq A$ be measurable, then $\lambda(A) = \lambda(B) + \lambda(A - B)$. As $|\lambda(A)| < \infty$ and λ can take at most one of the values ∞ or $-\infty$, so $|\lambda(B)| < \infty$.

(ii) It is a direct consequence of (i). □

Proposition 6.1.6 (Continuity of Signed Measure). *Let λ be a signed measure on (X,Σ) and $\{A_n\}$ a collection of members in Σ.*

(i) If $\{A_n\}$ is ascending, then

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

(ii) If $\{A_n\}$ is descending and one of $\lambda(A_n)$'s is finite, then

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

Proof. The proof is identical to Theorem 2.7.3. □

Definition 6.1.7. Let λ be a signed measure on (X,Σ), a set $E \in \Sigma$ is said to be **positive w.r.t. λ** if for every measurable $A \subseteq E$, $\lambda(A) \geq 0$, and said to be **negative w.r.t. λ** if for every measurable $A \subseteq E$, $\lambda(A) \leq 0$. Moreover, E is said to be **λ-null** if E is both positive and negative w.r.t. λ, i.e., every measurable subset of E has λ-measure zero.

We shall drop the reference "w.r.t. λ" in Definition 6.1.7 if the signed measure is unique and understood in the content.

Lemma 6.1.8. *Let λ be a signed measure on (X,Σ), then:*

(i) Any measurable subset of a positive set is positive.

(ii) Any countable union of positive sets is positive.

(iii) If E is measurable and $\lambda(E) \in (0,\infty)$, then there is a measurable set $A \subseteq E$ such that A is positive and $\lambda(A) > 0$.

Proof. (i) Let E be positive and $F \subseteq E$. To show F is positive, let $A \subseteq F$, then $A \subseteq E$, and hence $\lambda(A) \geq 0$ as E is positive.

(ii) Let E_n's be positive, let $F_1 = E_1$ and $F_n = E_n - E_1 - \cdots - E_{n-1}$ for $n \geq 2$, then $F_n \subseteq E_n$ are still positive by (i), and $\bigsqcup F_n = \bigcup E_n$. Let $A \subseteq \bigcup E_n$, then $A = \bigsqcup(A \cap F_n)$, so $\lambda(A) = \sum \lambda(A \cap F_n) \geq 0$.

(iii) If E is itself positive, then we are done. Otherwise there is a measurable $A_1 \subseteq E$ such that $\lambda(A_1) < 0$. We may find an $n \in \mathbb{N}$ such that $\lambda(A_1) < -\frac{1}{n}$, hence we may define

$$n_1 = \min\left\{n \in \mathbb{N} : \text{there is measurable } B \subseteq E,\ \lambda(B) < -\frac{1}{n}\right\}.$$

Let $E_1 \subseteq E$ be measurable such that $\lambda(E_1) < -\frac{1}{n_1}$. If $E - E_1$ is positive, then we are done. Otherwise there is a measurable $A_2 \subseteq E - E_1$ such that $\lambda(A_2) < 0$. Thus we can define

$$n_2 = \min\left\{n \in \mathbb{N} : \text{there is measurable } B \subseteq E - E_1,\ \lambda(B) < -\frac{1}{n}\right\}.$$

We let $E_2 \subseteq E - E_1$ such that $\lambda(E_2) < -\frac{1}{n_2}$. Inductively, if the procedure cannot terminate, we can define $n_k = \min\{n \in \mathbb{N} : \text{there is measurable } B \subseteq E - E_1 - \cdots - E_{k-1}, \lambda(B) < -\frac{1}{n}\}$ and let $E_k \subseteq E - E_1 - \cdots - E_{k-1}$ be measurable such that $\lambda(E_k) < -\frac{1}{n_k}$. Now E_k's are pairwise disjoint. Consider the equality

$$\lambda(E) = \sum_{k=1}^{\infty} \lambda(E_k) + \lambda\left(E - \bigsqcup_{k=1}^{\infty} E_k\right),$$

since $|\lambda(E)| < \infty$, $\sum_{k=1}^{\infty} \lambda(E_k)$ converges absolutely, so that $\lim_{k\to\infty} n_k = \infty$. Next since $\sum_{k=1}^{\infty} \lambda(E_k) < 0$, $\lambda(E - \bigsqcup_{k=1}^{\infty} E_k) > 0$ and $E - \bigsqcup_{k=1}^{\infty} E_k$ is positive. Indeed, let $A \subseteq E - \bigsqcup_{k=1}^{\infty} E_k$ be measurable, then $A \subseteq E - \bigsqcup_{k=1}^{p} E_k$ for each $p \in \mathbb{N}$, so $\lambda(A) \geq -\frac{1}{n_p - 1}$ for each large enough p, we conclude $\lambda(A) \geq 0$. $\square$

Remark. (i) and (ii) of Lemma 6.1.8 are also true if positive is replaced by negative.

Theorem 6.1.9 (Hahn Decomposition). *If λ is a signed measure on (X,Σ), there are a positive set P and a negative set N w.r.t. λ such that $P \cup N = X$ and $P \cap N = \emptyset$. If P', N' is another such pair, then $P \Delta P' = N \Delta N'$ are λ-null.*

Assume such decomposition does exist. Let A be any positive set. Then $\lambda(A) = \lambda(A \cap P) \leq \lambda(P)$, meaning that necessarily $\lambda(P) = \sup\{\lambda(A) : A \text{ is positive}\}$. By this observation, we try to construct a set P such that $\lambda(P)$ is the supremum of λ-measures of positive sets. Hopefully P is positive and its complement $X - P$ is negative.

Proof. WLOG we assume $\lambda(E) < \infty$ for each $E \in \Sigma$ (otherwise consider $-\lambda$). Let $m = \sup\{\lambda(A) : A \text{ is positive}\}$. Then there is a sequence of positive sets $\{P_n\}$ such that $m = \lim_{n\to\infty} \lambda(P_n)$. $P := \bigcup_{n=1}^\infty P_n$ is still positive and $\lambda(P) = m$, therefore $m < \infty$. Let $N = X - P$, our goal is to show N is negative.

Suppose N is not negative, then there is measurable $A \subseteq N$ such that $\infty > \lambda(A) > 0$. By (iii) of Lemma 6.1.8 there is $B \subseteq A$ such that B is positive and $\lambda(B) > 0$. But $m \geq \lambda(B \sqcup P) = \lambda(B) + m$. i.e., $\lambda(B) = 0$, a contradiction.

Finally if P' is positive and N' is negative such that $P' \cup N' = X$ and $P' \cap N' = \emptyset$, then $P \Delta P' = (P - P') \sqcup (P' - P) = \big(P \cap (X - P')\big) \sqcup \big(P' \cap (X - P)\big) = \big((X - N) \cap N'\big) \sqcup \big((X - N') \cap N\big) = (N' - N) \sqcup (N - N') = N \Delta N'$ is both positive and negative w.r.t. λ, hence λ-null. $\square$

Definition 6.1.10. The pair of subsets P,N of X in Theorem 6.1.9 is called a **Hahn decomposition for λ**.

6.1.2 Jordan Decompositions

Hahn decomposition is usually not unique as λ-null subset of P can be transferred to N, and vice versa. But it provides us with a *natural* and *unique* way to express λ as a difference of two positive measures. To state this result precisely, we need a new concept.

Definition 6.1.11. We say that two signed measures λ_1,λ_2 on (X,Σ) are **mutually singular** (or λ_1 is **singular with respect to** λ_2, or vice versa), denoted by

$$\lambda_1 \perp \lambda_2,$$

if there are measurable subsets E,F of X such that $E \sqcup F = X$, E is λ_2-null and F is λ_1-null.

If $\lambda(E) = 0$ whenever $E \cap A = \emptyset$, or equivalently $\lambda(E) = \lambda(E \cap A)$ for any $E \in \Sigma$, it is also common to say that **λ concentrates on A**[(1)]. So put in other way, $\lambda_1 \perp \lambda_2$ iff there are measurable E,F such that $E \sqcup F = X$, and λ_1,λ_2 concentrates on E,F respectively. For convenience we sometimes write symbolically $\lambda \to A$ to mean λ concentrates on A[(2)].

Theorem 6.1.12 (Jordan Decomposition). *If λ is a signed measure on (X,Σ), then there are unique positive measures λ^+,λ^- such that $\lambda = \lambda^+ - \lambda^-$ and $\lambda^+ \perp \lambda^-$.*

Proof. Let P,N be a Hahn decomposition for λ such that P is positive and N is negative. Set $\lambda^+(E) = \lambda(E \cap P)$ and $\lambda^-(E) = -\lambda(E \cap N)$, then both λ^+ and λ^- are positive measures. Now $\lambda = \lambda^+ - \lambda^-$, $\lambda^+ \to P$ and $\lambda^- \to N$, so that $\lambda^+ \perp \lambda^-$, thus we have established the existence part of the theorem.

[(1)]By using terminology in Chapter 7, A is a kind of "support" of λ.
[(2)]Warning! It is not a common notation.

Suppose there are positive measures μ and ν such that $\lambda = \mu - \nu$ and $\mu \perp \nu$. Let H,K be such that $H \sqcup K = X$, $\mu \to H$ and $\nu \to K$, then H,K forms another Hahn decomposition for λ[(3)]. By Theorem 6.1.9 $H\Delta P = K\Delta N$ are λ-null. Now $\mu(E) = \mu(E\cap H) = \lambda(E\cap H) + \nu(E\cap H) = \lambda(E\cap H) = \lambda(E\cap P) = \lambda^+(E)$ and similarly $\nu(E) = \nu(E\cap K) = \mu(E\cap K) - \lambda(E\cap K) = -\lambda(E\cap K) = -\lambda(E\cap N) = \lambda^-(E)$. □

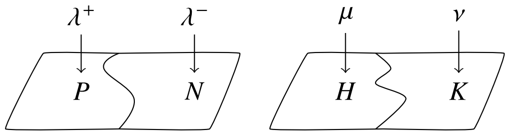


Figure 6.1: Different concentrations.

The positive measures λ^+ and λ^- are called the **positive** and **negative variations** of λ, whereas $\lambda = \lambda^+ - \lambda^-$ is called the **Jordan decomposition of $\boldsymbol{\lambda}$**. We further define the **total variation** of λ to be the measure

$$|\lambda| = \lambda^+ + \lambda^-.$$

Of course if λ is positive, then $|\lambda| = \lambda$.

In the study of measurable functions, the decomposition $f = f^+ - f^-$ enables us to break the proofs of various statements into two steps. One is for nonnegative measurable functions and one is for the general ones. Usually the second step is a direct consequence of the first step. The situation is similar due to Jordan decomposition theorem.

The name of λ^+, λ^- and $|\lambda|$ are due to the formulas given in the following theorem.

Theorem 6.1.13. *Let λ be a signed measure on (X,Σ), then for each $E \in \Sigma$:*

(i) $\lambda^+(E) = \sup\{\lambda(F) : F$ *is measurable subset of* $E\}$.

(ii) $\lambda^-(E) = \sup\{-\lambda(F) : F$ *is measurable subset of* $E\}$.

(iii) $$|\lambda|(E) = \sup\left\{\sum_{i=1}^{n} |\lambda(E_i)| : \{E_i\}_{i=1}^n \text{ is a measurable partition of } E, n \geq 1\right\}.$$

Proof. (i) Let P,N be a Hahn decomposition for λ, where P is positive and N is negative. Now for every $F \subseteq E$,

$$\lambda(F) = \lambda^+(F) - \lambda^-(F) \leq \lambda^+(F) \leq \lambda^+(E).$$

Since $E\cap P \subseteq E$ and $\lambda(E\cap P) =: \lambda^+(E)$, hence (i) follows.

(ii) It follows from the formula that $\lambda^- = (-\lambda)^+$.

(iii) Let P,N be defined as in (i), let $\{E_i\}_{i=1}^n$ be a measurable partition of E,

$$\sum_{i=1}^{n} |\lambda(E_i)| \leq \sum_{i=1}^{n} |\lambda|(E_i) = |\lambda|(E).$$

(3) We indicate different concentrations in Figure 6.1.

Consider the measurable partition $\{E \cap P, E \cap N\}$, one has

$$|\lambda(E \cap P)| + |\lambda(E \cap N)| = \lambda^+(E) + \lambda^-(E) = |\lambda|(E),$$

hence (iii) follows. □

Definition 6.1.14. Let λ be a signed measure and μ a positive measure on (X,Σ). We say that λ is **absolutely continuous with respect to μ**, denoted by $\lambda \ll \mu$, if for every measurable A,

$$\mu(A) = 0 \implies \lambda(A) = 0.$$

The name used in Definition 6.1.14 for signed measures comes from the following result for finite signed measures.

Theorem 6.1.15. *Let λ be a finite signed measure and μ a positive measure on (X,Σ), then $\lambda \ll \mu$ iff for every $\epsilon > 0$, there is $\delta > 0$ such that $\mu(E) < \delta \implies |\lambda(E)| < \epsilon$.*

For simplicity we refer the latter condition to "ϵ-δ condition". The proof requires a simple observation that given a signed measure λ and a positive measure μ, $\lambda \ll \mu$ iff $|\lambda| \ll \mu$ (detail can be found in part (v) of Proposition 6.1.16)).

Proof. Suppose the ϵ-δ condition holds, let E be measurable such that $\mu(E) = 0$, then $\mu(E) < \delta$ for any $\delta > 0$, so that $|\lambda(E)| < \epsilon$ for any $\epsilon > 0$, $\lambda(E) = 0$.

Conversely, assume $\lambda \ll \mu$. Suppose the ϵ-δ condition fails, then there is an $\epsilon > 0$ such that for each $n \in \mathbb{N}$, there is $E_n \in \Sigma$ such that $\mu(E_n) < \frac{1}{2^n}$ and $|\lambda(E_n)| \geq \epsilon$. The latter inequality implies that $|\lambda|(E_n) \geq \epsilon$, so for each N, $|\lambda|(\bigcup_{n=N}^\infty E_n) \geq \epsilon$ and $\mu(\bigcup_{n=N}^\infty E_n) \leq \frac{1}{2^N}$, which is a contradiction as $A := \lim_{N\to\infty} \bigcup_{n=N}^\infty E_n$ has μ-measure zero, but $|\lambda|(A) \geq \epsilon$. □

Remark. Let μ be a positive measure on (X,Σ) and let $f : X \to [-\infty,\infty]$ be integrable, then a finite signed measure λ on Σ defined by $\lambda(E) = \int_E f\,d\mu$ is absolutely continuous w.r.t. μ, hence it satisfies the ϵ-δ condition. i.e., for any $\epsilon > 0$, we can find a $\delta > 0$ such that $\mu(E) < \delta \implies \left|\int_E f\,d\mu\right| < \epsilon$. Which is exactly the result in Problem 5.12.

Proposition 6.1.16. *Let μ be a positive measure and λ, λ_1 and λ_2 signed measures on (X,Σ), then:*

(i) If $\lambda \to A$, then $|\lambda| \to A$.

(ii) If $\lambda_1 \perp \lambda_2$, then $|\lambda_1| \perp |\lambda_2|$.

(iii) If $\lambda_1 \perp \mu$ and $\lambda_2 \perp \mu$, then for each $c \in \mathbb{R}$, $c\lambda_1 \perp \mu$ and $\lambda_1 + \lambda_2 \perp \mu$ if $\lambda_1 + \lambda_2$ is also a signed measure.

(iv) If $\lambda_1 \ll \mu$ and $\lambda_2 \ll \mu$, then for each $c \in \mathbb{R}$, $c\lambda_1 \ll \mu$ and $\lambda_1 + \lambda_2 \ll \mu$ if $\lambda_1 + \lambda_2$ is also a signed measure.

(v) $\lambda \ll \mu$ iff $|\lambda| \ll \mu$.

(vi) If $\lambda_1 \ll \mu$ and $\lambda_2 \perp \mu$, then $\lambda_1 \perp \lambda_2$.

(vii) If $\lambda \ll \mu$ and $\lambda \perp \mu$, then $\lambda = 0$.

Proof. (i) Let $E \subseteq X - A$, and let $\{E_i\}_{i=1}^n$ be a measurable partition of E, then $E_i \subseteq X - A$. As $\lambda \to A$, $\lambda(E_i) = 0$ for each i. Hence $\sum_{i=1}^n |\lambda(E_i)| = 0$. By formula in Definition 6.1.14, $|\lambda|(E) = 0$, so $|\lambda| \to A$.

(ii) If $\lambda_1 \perp \lambda_2$, then there are measurable A, B such that $A \sqcup B = X$ with $\lambda_1 \to A$ and $\lambda_2 \to B$. By (i), $|\lambda_1| \to A$ and $|\lambda_2| \to B$, so $|\lambda_1| \perp |\lambda_2|$.

(iii) Let A_i, B_i be such that for $i = 1,2$, $A_i \sqcup B_i = X$, $\lambda_i \to A_i$ and $\mu \to B_i$. To show $c\lambda_1 \perp \mu$, it suffices to show $c\lambda_1 \to A_1$. Let $E \subseteq X - A_1$, then $\lambda_1(E) = 0$, so $c\lambda_1(E) = 0$, as desired. Next, $\lambda_1 + \lambda_2$ concentrates on $A_1 \cup A_2$ because if $E \subseteq X - A_1 - A_2$, $\lambda_1(E) = \lambda_2(E) = 0$. We need to show μ concentrates on $X - A_1 \cup A_2 = B_1 \cap B_2$, which is obvious.

(iv) It is obvious.

(v) Assume $\mu(E) = 0$, let $\{E_i\}_{i=1}^n$ be a measurable partition of E, then $\mu(E_i) = 0$ so that $\lambda(E_i) = 0$ for each i and $\sum_{i=1}^n |\lambda(E_i)| = 0$, we conclude $|\lambda|(E) = 0$, so $|\lambda| \ll \mu$. The converse if obvious since $|\lambda(E)| \le |\lambda|(E)$.

(vi) Let A, B be measurable such that $A \sqcup B = X$, $\lambda_2 \to A$ and $\mu \to B$. It is enough to show $\lambda \to B$, that is because for measurable $E \subseteq X - B$, $\mu(E) = 0$, so by absolute continuity $\lambda_1(E) = 0$.

(vii) Let A, B be measurable such that $A \sqcup B = X$ and $\lambda \to A$ and $\mu \to B$. By the proof of (vi), $\lambda \to B$, so $\lambda = 0$. □

6.1.3 Lebesgue-Radon-Nikodym Theorem

Lemma 6.1.17. *If μ is a σ-finite positive measure on (X, Σ), then there is a $w \in L^1(X, \mu)$ such that $0 < w(x) < 1$ for each $x \in X$.*

Proof. Let $X = \bigsqcup_{n=1}^\infty X_n$ with $\mu(X_n) < \infty$. For each n we define a function $w_n : X \to \mathbb{R}$ by

$$w_n(x) = \begin{cases} \dfrac{1}{2^n} \cdot \dfrac{1}{1 + \mu(X_n)}, & x \in X_n, \\ 0, & x \in X - X_n. \end{cases}$$

Define $w = \sum_{n=1}^\infty w_n$, then $0 < w(x) < 1$ for each $x \in X$ and $\int_X w\, d\mu = \sum_{n=1}^\infty \int_{X_n} w_n\, d\mu = \sum_{n=1}^\infty \frac{\mu(X_n)}{2^n(1+\mu(X_n))} < 1$. □

By Lemma 6.1.17 every σ-finite positive measure μ on X induces a finite measure $w\, d\mu$[(4)] which has the same collection of sets of measure zero with μ since $w(x) > 0$ for each $x \in X$. i.e., $\int_E w\, d\mu = 0$ iff $\mu(E) = 0$.

To describe the general decomposition of a signed measure, we need to introduce a larger class of "integrable" functions.

Definition 6.1.18. A measurable function $f : X \to [-\infty, \infty]$ is said to be **extended μ-integrable** if at least one of $\int_X f^+\, d\mu$ and $\int_X f^-\, d\mu$ is finite. In this case, we define

$$\int_E f\, d\mu = \int_E f^+\, d\mu - \int_E f^-\, d\mu$$

as before.

[(4)]It is a convention to denote $d\nu := f\, d\mu$ a set function defined by $\nu(E) = \int_E f\, d\mu$. Note that it does not mean $\int_E g\, d\nu = \int_E gf\, d\mu$, to have this kind of equality we have to be careful.

Example 6.1.4 tells us $\lambda(E) := \int_E f\,d\mu$ is a signed measure if f is extended μ-integrable. In fact integration against such functions is a rich source of signed measures:

Theorem 6.1.19 (Lebesgue-Radon-Nikodym). *Let μ be a σ-finite positive measure and λ a σ-finite signed measure on (X,Σ), then:*

(i) There is a unique pair of signed measures λ_a, λ_s on Σ such that

$$\lambda = \lambda_a + \lambda_s, \quad \text{where } \lambda_a \ll \mu \text{ and } \lambda_s \perp \mu.$$

If λ is positive (and finite), λ_a and λ_s are also positive (and finite).

(ii) There is an $h : X \to [-\infty,\infty]$ that is extended μ-integrable such that $d\lambda_a = f\,d\mu$. If λ is positive (and finite), then h is nonnegative (and integrable). If h' is another such function, then $h = h'$ μ-a.e..

The pair λ_a,λ_s is unique since if λ'_a,λ'_s is another such pair, then the equality $\lambda_a - \lambda'_a = \lambda'_s - \lambda_s$ implies $\lambda_a - \lambda'_a \ll \mu$ and $\lambda_a - \lambda'_a \perp \mu$, so $\lambda_a = \lambda'_a$, and thus $\lambda_s = \lambda'_s$. To be more careful we need some σ-finite argument to avoid $\infty - \infty$ case. We leave all specific checkings for exercises.

Proof. We will prove (i) and (ii) at the same time. Assume first that λ is a finite positive measure. Since μ is σ-finite, by Lemma 6.1.17 there is a function $w \in L^1(X,\mu)$ such that $0 < w(x) < 1$ for each $x \in X$. The measure $d\varphi := d\lambda + w\,d\mu$ is finite and positive, $\lambda \le \varphi$. For $f \in L^1(\varphi)$,

$$\left|\int_X f\,d\lambda\right| \le \int_X |f|\,d\varphi \le \left(\int_X |f|^2\,d\varphi\right)^{1/2} (\varphi(X))^{1/2},$$

so $f \mapsto \int_X f\,d\lambda$ is a bounded linear functional on the Hilbert space $L^2(\varphi)$, thus this functional must be given by an inner product, i.e., there is a $g \in L^2(\varphi)$ such that for each $f \in L^2(\varphi)$,

$$\int_X f\,d\lambda = \int_X fg\,d\varphi. \tag{6.1.20}$$

For each $E \in \Sigma$ and $\varphi(E) > 0$, we may set $f = \chi_E$ in (6.1.20) to obtain

$$\frac{1}{\varphi(E)}\int_E g\,d\varphi = \frac{\lambda(E)}{\varphi(E)} \in [0,1],$$

hence by Theorem 5.2.43, $0 \le g \le 1$ φ-a.e. on X. Without affecting (6.1.20), we may redefine g if necessary on a set of φ-measure zero, so let's assume $0 \le g \le 1$ on X. As w is nonnegative[5], we can rewrite (6.1.20) as

$$\int_X f(1-g)\,d\lambda = \int_X fgw\,d\mu. \tag{6.1.21}$$

Define

$$A = \{x \in X : 0 \le g(x) < 1\}, \quad B = \{x \in X : g(x) = 1\}.$$

[5] If μ and ν are positive, then for every $f \ge 0$ or $f \in L^1(\mu+\nu)$, $\int_X f\,d(\mu+\nu) = \int_X f\,d\mu + \int_X f\,d\nu$. Also if $d\lambda = g\,d\mu$, for some $g \ge 0$, then for every $f \ge 0$ or $f \in L^1(\lambda)$, $\int_X f\,d\lambda = \int_X fg\,d\mu$.

By setting $f = \chi_B$, $\mu(B) = 0$, so $\mu \to X - B$, if we define $\lambda_s(E) = \lambda(E \cap B)$, then $\lambda_s \to B$ and hence $\mu \perp \lambda_s$. For $E \in \Sigma$, define $\lambda_a(E) = \lambda(E \cap A)$ and let $f = (1 + g + \cdots + g^n)\chi_E$ in (6.1.21),

$$\int_{E\cap A}(1 - g^{n+1})\,d\lambda = \int_E (g + g^2 + \cdots + g^{n+1})w\,d\mu.$$

$1 - g^{n+1} \nearrow 1$ on $E \cap A$ and $(g + g^2 + \cdots + g^{n+1})w \nearrow h \geq 0$ on E, where h is measurable, hence by monotone convergence theorem,

$$\lambda_a(E) := \lambda(E \cap A) = \int_E h\,d\mu.$$

Since λ is a finite positive measure, so h is nonnegative and integrable.

If λ is σ-finite and positive, let X_n's be disjoint such that $X = \bigsqcup_{n=1}^{\infty} X_n$ and $\lambda(X_n) < \infty$. Now $\lambda(E) = \sum_{n=1}^{\infty} \lambda(E \cap X_n)$. Define $\lambda_n(E) = \lambda(E \cap X_n)$ for each $E \in \Sigma$, then λ_n is a finite positive measure on (X,Σ), by the preceding case there are a measurable $h_n : X \to [0,\infty)$ and a positive measure $\lambda_s^n \perp \mu$ on Σ such that $\lambda_n = h_n\,d\mu + d\lambda_s^n$. Hence

$$\lambda(E) = \sum_{n=1}^{\infty} \lambda_n(E \cap X_n) = \sum_{n=1}^{\infty} \int_{E\cap X_n} h_n\,d\mu + \sum_{n=1}^{\infty} \lambda_s^n(E \cap X_n). \tag{6.1.22}$$

Let $h = \sum_{n=1}^{\infty} h_n \chi_{X_n}$ and $\lambda_s(E) = \sum_{n=1}^{\infty} \lambda_s^n(E \cap X_n)$, then λ_s is a positive measure that is singular w.r.t. μ, (6.1.22) becomes

$$\lambda(E) = \sum_{n=1}^{\infty} \int_E h_n \chi_{X_n}\,d\mu + \lambda_s(E) = \int_E h\,d\mu + \lambda_s(E).$$

Note that this time h may not be integrable.

Finally if λ is a σ-finite signed measure, then consider the Jordan decomposition of λ, i.e., $\lambda = \lambda^+ - \lambda^-$. Again we apply the preceding case to λ^+ and λ^-. As one of λ^+ and λ^- must be finite, so the resulting h is extended μ-integrable whose uniqueness is left as an exercise. □

Remark. Actually some effort has to be paid to finish the proof of Theorem 6.1.19. WLOG, assume λ^- is finite. After writing $\lambda^+ = h_1\,d\mu + \lambda_s$ and $\lambda^- = h_2\,d\mu + \lambda_s'$, then of course $\lambda_s - \lambda_s'$ is a signed measure. What we are concerned about is if $h_1\,d\mu - h_2\,d\mu$ can be expressed as $h\,d\mu$, for some extended μ-integrable h. The canonical choice is to choose $h = h_1 - h_2$, for this to make sense we can further assume $h_2 \neq \infty$ on X.

We need to check $(h_1 - h_2)\,d\mu$ is indeed a signed measure, after that we conclude

$$h_1\,d\mu - h_2\,d\mu = (h_1 - h_2)\,d\mu \tag{6.1.23}$$

due to σ-finiteness of λ. To be specific, the equality (6.1.23) holds for measurable set that has finite λ-measure. Suppose in the worst case $\int_X h_1\,d\mu = +\infty$ (h_2 is integrable since λ^- is finite), we check that $h_1 - h_2$ is extended μ-integrable by showing that its negative parts is μ-integrable, which does because $(h_1 - h_2)^- = (-(h_1 - h_2)) \vee 0 = \frac{1}{2}(-(h_1 - h_2) + |h_1 - h_2|) \leq h_2$. (6.1.23) now follows from continuity of measure. Uniqueness of h also follows from σ-finite argument.

The decomposition $\lambda = \lambda_a + \lambda_s$, where $\lambda_a \ll \mu$ and $\lambda \perp \mu$ is called the **Lebesgue decomposition of $\boldsymbol{\lambda}$ w.r.t. $\boldsymbol{\mu}$**. Which is unique by the remark preceding the proof of

Theorem 6.1.19. Special case of Theorem 6.1.19 is of our particular interest. Suppose $\lambda \ll \mu$, where λ is signed and μ is positive and both are σ-finite, then $d\lambda = d\lambda_a = f\,d\mu$, for some extended μ-integrable function $f : X \to [-\infty,\infty]$. The "uniquely" determined f is denoted by $d\lambda/d\mu$, i.e.,

$$d\lambda = \frac{d\lambda}{d\mu}\,d\mu.$$

This result is known as Radon-Nikodym theorem and $d\lambda/d\mu$ is called the **Radon-Nikodym derivative**.

Let λ be a signed measure on (X,Σ) and write $\lambda = \lambda^+ - \lambda^-$, then for any measurable set E, let $f = \chi_E$, one has

$$\int_X f\,d\lambda = \int_X f\,d\lambda^+ - \int_X f\,d\lambda^-.$$

This formula suggests a (only) reasonable way to define integral of measurable $f : X \to [-\infty,\infty]$ w.r.t. a signed measure.

Definition 6.1.24. Let λ be a signed measure on (X,Σ), a function $f : X \to [-\infty,\infty]$ is said to be integrable w.r.t. λ if f is integrable w.r.t. both λ^+ and λ^-. In this case we define $L^1(X,\lambda) = L^1(X,\lambda^+) \cap L^1(X,\lambda^-)$ and the integral of f w.r.t. λ over $E \in \Sigma$ is

$$\int_E f\,d\lambda = \int_E f\,d\lambda^+ - \int_E f\,d\lambda^-.$$

Proposition 6.1.25. *Let λ be a σ-finite signed measure and let ν,μ be σ-finite positive measures on (X,Σ) such that $\lambda \ll \nu$ and $\nu \ll \mu$.*

(i) If $f \in L^1(X,\lambda)$, then $f \cdot \dfrac{d\lambda}{d\nu} \in L^1(X,\nu)$ and $\displaystyle\int_X f\,d\lambda = \int_X f \cdot \frac{d\lambda}{d\nu}\,d\nu$.

(ii) $\lambda \ll \mu$ and $\dfrac{d\lambda}{d\mu} = \dfrac{d\lambda}{d\nu} \cdot \dfrac{d\nu}{d\mu}$ μ-a.e..

Proof. (i) We first assume λ is positive. Let E be measurable and put $f = \chi_E$, then the equality $d\lambda = (d\lambda/d\nu)\,d\nu$ implies

$$\int_X f\,d\lambda = \int_X f \cdot \frac{d\lambda}{d\nu}\,d\nu. \tag{6.1.26}$$

By linearity (6.1.26) holds when f is replaced by nonnegative simple functions, and hence it holds by monotone convergence theorem when f is replaced by nonnegative measurable functions. If f is general integrable function, then we are done by considering $f = f^+ - f^-$ ($d\lambda/d\nu$ is nonnegative).

In general if λ is signed, then $\lambda = \lambda^+ - \lambda^-$. Since f is integrable w.r.t. both λ^+ and λ^-, we can write

$$\int_X f\,d\lambda = \int_X f\,d\lambda^+ - \int_X f\,d\lambda^- = \int_X f\frac{d\lambda^+}{d\nu}\,d\nu - \int_X f\frac{d\lambda^-}{d\nu}\,d\nu = \int_X f\frac{d\lambda}{d\nu}\,d\nu.$$

Here we have used the facts that $\lambda \ll \nu$ iff $|\lambda| \ll \nu$ iff $\lambda^+,\lambda^- \ll \nu$ and also that $\frac{d\lambda}{d\nu} = \frac{d\lambda^+}{d\nu} - \frac{d\lambda^-}{d\nu}$ ν-a.e.[(6)].

[(6)]This is because $\int_E \frac{d\lambda}{d\nu}\,d\nu = \lambda(E) = \lambda^+(E) - \lambda^-(E) = \int_E \frac{d\lambda^+}{d\nu}\,d\nu - \int_E \frac{d\lambda^-}{d\nu}\,d\nu = \int_E (\frac{d\lambda^+}{d\nu} - \frac{d\lambda^-}{d\nu})\,d\nu$, for every E with finite λ-measure, and hence every E measurable due to σ-finiteness of λ, so $\frac{d\lambda}{d\nu} = \frac{d\lambda^+}{d\nu} - \frac{d\lambda^-}{d\nu}$ ν-a.e..

(ii) Let $\mu(E) = 0$, then $\nu(E) = 0$ and hence $\lambda(E) = 0$, meaning that $\lambda \ll \mu$. For any E measurable, $\int_E \frac{d\lambda}{d\mu}\, d\mu = \lambda(E) = \int_E \frac{d\lambda}{d\nu}\, d\nu = \int_E \frac{d\lambda}{d\nu} \cdot \frac{d\nu}{d\mu}\, d\mu$. $\square$

6.2 Complex Measures

6.2.1 Total Variation Measure

Definition 6.2.1. Let (X,Σ) be a measurable space, a set function $\lambda : \Sigma \to \mathbb{C}$ is called a complex measure if it has the following properties:

(i) $\lambda(\emptyset) = 0$.

(ii) If $\{E_n\}_{n=1}^{\infty}$ is a disjoint collection of members in Σ, then

$$\lambda\left(\bigsqcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \lambda(E_n). \tag{6.2.2}$$

As in the earlier issues discussed for signed measure, the convergence in (6.2.2) by definition is absolute. Define $\lambda_r(E) = \operatorname{Re}\lambda(E)$ and $\lambda_i(E) = \operatorname{Im}\lambda(E)$, it is obvious that λ_r and λ_i are finite signed measures on Σ.

We now introduce a positive measure induced by a complex measure λ whose construction is motivated by the following problem: We need to find the smallest positive measure μ that dominates λ in the sense that for every measurable set E in (X,Σ), $\mu(E) \geq |\lambda(E)|$. Let $\{E_i\}_{i=1}^n$ be a measurable partition of E, one has $\mu(E) \geq \sum_{i=1}^n |\lambda(E_i)|$, so that

$$\mu(E) \geq |\lambda|(E) := \sup \sum_{i=1}^{n} |\lambda(E_i)|. \tag{6.2.3}$$

Where the supremum on the RHS of (6.2.3) ranges over all partitions of E into finite disjoint measurable subsets. Inspired by (iii) of theorem 6.1.13 $|\lambda|$ is expected to be a positive measure on Σ (which is indeed the case when λ is a finite signed measure). As an analogue to signed measures, $|\lambda|$ is also called the **total variation of $\boldsymbol{\lambda}$**. It turns out that $|\lambda|(X) < \infty$, which we shall prove shortly, and hence we say that λ is of **bounded variation**.

Of course to make things complicated we can also redefine $|\lambda|$ in (6.2.3) by letting $n = \infty$ and the supremum this time ranges over all partitions of E into countably many disjoint measurable subsets. Both definitions are commonly used, meaning that in fact these two set functions are indeed the same.

To make this precise, for $E \in \Sigma$ we denote $\pi_{<\infty}(E)$ the collection of all finite collection $\{E_i\}_{i=1}^n$ which forms a measurable partition of E. We also denote likewise $\pi_\infty(E)$ the collection of all measurable partition of E, each partition consists of countably many disjoint measurable subsets of E.

Proposition 6.2.4. *Let E be a measurable subset of (X,Σ), then*

$$\begin{aligned} \mu(E) &:= \sup\left\{\sum |\lambda(E_i)| : \{E_i\} \in \pi_{<\infty}(E)\right\} \\ &= \sup\left\{\sum |\lambda(E_i)| : \{E_i\} \in \pi_{\infty}(E)\right\} =: \nu(E). \end{aligned}$$

Proof. Since $\pi_{<\infty}(E) \subseteq \pi_\infty(E)$, $\mu(E) \le \nu(E)$. To show $\nu(E) \le \mu(E)$, let $t \in \mathbb{R}$ be such that $t < \nu(E)$, then there is $\{E_i\}_{i=1}^\infty \in \pi_\infty(E)$ such that $t < \sum_{i=1}^\infty |\lambda(E_i)|$, so there is an n such that $t < \sum_{i=1}^n |\lambda(E_i)|$. Since $\{E_1,\dots,E_n, E - \bigsqcup_{i=1}^n E_i\} \in \pi_{<\infty}(E)$, one has

$$t < \sum_{i=1}^n |\lambda(E_i)| \le \sum_{i=1}^n |\lambda(E_i)| + \left|\lambda\left(E - \bigsqcup_{i=1}^n E_i\right)\right| \le \mu(E).$$

This is true for each $t < \nu(E)$, we conclude $\nu(E) \le \mu(E)$. □

We are free to use any one of the definitions. For convenience we will not be so specific whether a partition is finite or countably infinite.

Theorem 6.2.5. *The total variation $|\lambda|$ of a complex measure λ on (X,Σ) is a positive measure on Σ.*

Proof. Let E_i's be disjoint and measurable, write $E = \bigsqcup_i E_i$. For each i, choose a $t_i < |\lambda|(E_i)$, then there is a measurable partition $\{F_{ij}\}$ of E_i such that $t_i < \sum_j |\lambda(F_{ij})|$. So

$$\sum_i t_i \le \sum_i \sum_j |\lambda(F_{ij})| \le |\lambda|(E).$$

We can choose t_i as closed to $|\lambda|(E_i)$ as we want, $\sum_i |\lambda|(E_i) \le |\lambda|(E)$.

To prove the reverse inequality, let $\{A_j\}$ be any measurable partition of E, then

$$\sum_j |\lambda(A_j)| = \sum_j \left|\sum_i \lambda(A_j \cap E_i)\right| \le \sum_i \sum_j |\lambda(A_j \cap E_i)| \le \sum_i |\lambda|(E_i).$$

We take supremum on LHS to obtain $|\lambda|(E) \le \sum_i |\lambda|(E_i)$. So $|\lambda|$ is countably additive. It is obvious that $|\lambda|(\emptyset) = 0$. □

Lemma 6.2.6. *For any $z_1, z_2, \dots, z_n \in \mathbb{C}$, there is a nonempty subset S of $\{1,2,\dots,n\}$ such that*

$$\left|\sum_{k\in S} z_k\right| \ge \frac{1}{\pi}\sum_{k=1}^n |z_k|.$$

Proof. Let $S \ne \emptyset$ be any subset of $\{1,2,\dots,n\}$. Write $z_k = |z_k| e^{i\alpha_k}$, then for any $\theta \in \mathbb{R}$,

$$\left|\sum_{k\in S} z_k\right| = \left|\sum_{k\in S} z_k e^{-i\theta}\right| \ge \left|\operatorname{Re}\sum_{k\in S} z_k e^{-i\theta}\right| = \left|\sum_{k\in S} |z_k|\cos(\alpha_k - \theta)\right|. \tag{6.2.7}$$

For a fixed θ, we can choose $S = S(\theta)$ to be the indexes of those α_k's such that $\cos(\alpha_k - \theta) > 0$ ($S(\theta)$ can be empty), then (6.2.7) becomes

$$\left|\sum_{k\in S(\theta)} z_k\right| \ge \sum_{k\in S(\theta)} |z_k|\cos(\alpha_k - \theta) = \sum_{k=1}^n |z_k|\cos^+(\alpha_k - \theta).$$

RHS is a continuous function in θ, we may choose $\theta = \theta_0$ such that RHS attains its maximum, then for any $\theta \in \mathbb{R}$,

$$\left|\sum_{k\in S(\theta_0)} z_k\right| \ge \sum_{k=1}^n |z_k|\cos^+(\alpha_k - \theta),$$

the result follows from integrating both sides over $[0,2\pi]$. □

Theorem 6.2.8. *If λ is a complex measure on (X,Σ), then $|\lambda|(X) < \infty$.*

Proof. Assume on the contrary that $|\lambda|(X) = \infty$. Then for any $N > 0$, there is a measurable partition $\{A_i\}_{i=1}^n$ of X such that $\sum_{i=1}^n |\lambda(A_i)| > N$. By Lemma 6.2.6 there is an $A \in \Sigma$ (a union of some A_i's) such that

$$|\lambda(A)| > \frac{N}{\pi}.$$

Let $B = X - A$, then $|\lambda(B)| = |\lambda(X) - \lambda(A)| \geq |\lambda(A)| - |\lambda(X)| > \frac{N}{\pi} - |\lambda(X)|$. We choose an N large at the beginning such that $|\lambda(A)|,|\lambda(B)| > 1$. As $|\lambda|$ is a positive measure, at least one of A and B must have ∞ $|\lambda|$-measure, say $|\lambda|(B) = \infty$.

Let $A_1 = A$ and $B_1 = B$. Repeat the same procedure to B, we can find disjoint $A_2,B_2 \subseteq B_1$ such that $|\lambda(A_2)| > 1$ and $|\lambda|(B_2) = \infty$, inductively, we can find disjoint $A_k,B_k \subseteq B_{k-1}$ such that $|\lambda(A_k)| > 1$ and $|\lambda|(B_k) = \infty$. Finally

$$\lambda\left(\bigsqcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \lambda(A_k)$$

and RHS does not converge absolutely, a contradiction. □

Given a complex Borel measure on X, we define

$$\|\lambda\| = |\lambda|(X).$$

As indicated by the notation, $\|\cdot\|$ defines a norm on the collection of complex Borel measures $M(X)$ on X. Not only that, $(M(X),\|\cdot\|)$ is a Banach space.

Now several concepts and results for signed measures can be directly translated to complex measures.

Definition 6.2.9. We say that two measures (can be signed or complex) λ_1,λ_2 on (X,Σ) **are mutually singular** (or λ_1 is **singular with respect to** λ_2, or vice versa), denoted by

$$\lambda_1 \perp \lambda_2,$$

if there are measurable subsets E,F of X such that λ_1 concentrates on E and λ_2 concentrates on F.

Definition 6.2.10. Let λ be a complex measure and μ a positive measure on (X,Σ). We say that λ is **absolutely continuous with respect to μ**, denoted by $\lambda \ll \mu$, if for every measurable A,

$$\mu(A) = 0 \implies \lambda(A) = 0.$$

Proposition 6.2.11. *Let μ be a positive measure and λ,λ_1 and λ_2 complex measures on (X,Σ), then:*

(i) If $\lambda \to A$, then $|\lambda| \to A$.

(ii) If $\lambda_1 \perp \lambda_2$, then $|\lambda_1| \perp |\lambda_2|$.

(iii) *If $\lambda_1 \perp \mu$ and $\lambda_2 \perp \mu$, then for each $c \in \mathbb{C}$, $c\lambda_1 \perp \mu$ and $\lambda_1 + \lambda_2 \perp \mu$.*

(iv) *If $\lambda_1 \ll \mu$ and $\lambda_2 \ll \mu$, then for each $c \in \mathbb{C}$, $c\lambda_1 \ll \mu$ and $\lambda_1 + \lambda_2 \ll \mu$.*

(v) *$\lambda \ll \mu$ iff $|\lambda| \ll \mu$.*

(vi) *If $\lambda_1 \ll \mu$ and $\lambda_2 \perp \mu$, then $\lambda_1 \perp \lambda_2$.*

(vii) *If $\lambda \ll \mu$ and $\lambda \perp \mu$, then $\lambda = 0$.*

The proof is, word by word, same as before.

6.2.2 Lebesgue-Radon-Nikodym Theorem

Theorem 6.2.12 (Lebesgue-Radon-Nikodym). *Let μ be a σ-finite positive measure on (X,Σ) and λ a complex measure on Σ.*

(i) *There is a unique pair of complex measures λ_a and λ_s on Σ such that*

$$\lambda = \lambda_a + \lambda_s, \quad \text{where } \lambda_a \ll \mu \text{ and } \lambda_s \perp \mu.$$

(ii) *There is a unique $h \in L^1(X,\mu)$ such that*

$$\lambda_a(E) = \int_E h\, d\mu$$

for every $E \in \Sigma$.

The uniqueness part of both measures and functions are essentially the same as Theorem 6.1.19, but a bit easier this time because everything is finite.

Proof. Write $\lambda = \lambda_r + i\lambda_i$. Then λ_r and λ_i are finite signed measures on Σ, hence by Theorem 6.1.19 there are extended real-valued integrable functions $h_r, h_i \geq 0$ and finite signed measures λ'_s and λ''_s that are singular w.r.t. μ such that

$$d\lambda_r = h_r\, d\mu + d\lambda'_s \quad \text{and} \quad d\lambda_i = h_i\, d\mu + d\lambda''_s.$$

Define $h = h_r + ih_i$ and $\lambda_s = \lambda'_s + \lambda''_s \perp \mu$,

$$d\lambda = d\lambda_r + id\lambda_i = (h_r + ih_i)\, d\mu + d\lambda'_s + id\lambda''_s = h\, d\mu + d\lambda_s. \qquad \square$$

6.2.3 Polar Representation

Theorem 6.2.13 (Polar Representation). *Let λ be a complex measure on (X,Σ), then there is a complex measurable function $h \in L^1(|\lambda|)$ such that $|h(x)| = 1$ for each $x \in X$ and*

$$d\lambda = h\, d|\lambda|.$$

Proof. It is obvious that $\lambda \ll |\lambda|$, by Radon-Nikodym theorem there is an $h \in L^1(|\lambda|)$ such that $d\lambda = h\, d|\lambda|$. Let $A_r = \{x \in X : |h(x)| < r\}$ and let $\{E_i\}$ be a measurable partition of A_r, then

$$\sum |\lambda(E_i)| \leq \sum \left| \int_{E_i} h\, d|\lambda| \right| \leq \sum r|\lambda|(E_i) = r|\lambda|(A_r),$$

which implies $|\lambda|(A_r) \leq r|\lambda|(A_r)$. Hence if $r < 1$, $|\lambda|(A_r) = 0$, which implies that $|h| \geq 1$ $|\lambda|$-a.e..

Let $E \in \Sigma$ and $|\lambda|(E) > 0$, then

$$\left| \frac{1}{|\lambda|(E)} \int_E h\, d|\lambda| \right| = \frac{|\lambda(E)|}{|\lambda|(E)} \leq 1,$$

hence $|h| \leq 1$ $|\lambda|$-a.e., so that $|h| = 1$ a.e.. Without affecting the equality $d\lambda = h\, d|\lambda|$, we may assume $|h(x)| = 1$ for each $x \in X$. □

Theorem 6.2.14. *Let μ be a positive measure on (X,Σ), $g \in L^1(\mu)$ and define $d\lambda = g\, d\mu$, then $d|\lambda| = |g|\, d\mu$.*

Proof. By Theorem 6.2.13 there is $h \in L^1(|\lambda|)$ such that $|h| = 1$ on X and $d\lambda = h\, d|\lambda|$. By hypothesis,

$$h\, d|\lambda| = g\, d\mu,$$

hence $d|\lambda| = \overline{h} g\, d\mu$ (why?). Since $\int_E \overline{h} g\, d\mu \geq 0$ for every $E \in \Sigma$, $\overline{h} g \geq 0$ μ-a.e., so $\overline{h} g = |g|$ μ-a.e., i.e., $d|\lambda| = |g|\, d\mu$. □

Remark. For $g \in L^1(\mu)$, where μ is a positive measure on X, we may also write $|g\, d\mu| = |g|\, d\mu$ if we accept the notation $\lambda = d\lambda$ and bear in mind that the only way to interpret the notation $(f\, d\mu)(E)$ is $\int_E f\, d\mu$.

Equality in Theorem 6.2.13 provides us with a reasonable way to define integral w.r.t. complex measures:

Definition 6.2.15. Let λ be a complex Borel measure on (X,Σ) and $d\lambda = h\, d|\lambda|$, where h is complex measurable and $|h| = 1$. We say that f is integrable w.r.t. λ if it does w.r.t. $|\lambda|$, and it that case we define the integral of f w.r.t. λ over $E \in \Sigma$ by

$$\int_E f\, d\lambda = \int_E f h\, d|\lambda|.$$

6.2.4 Bounded Linear Functionals on L^p

As an application of Radon-Nikodym theorem we try to identify the dual space of $L^p(X,\mu)$ when μ is σ-finite, positive and $1 \leq p < \infty$. It is one of the most concrete examples in the study of functional analysis. When $X = \mathbb{N}$ and $\mu = c$, the counting measure,

$$\ell^p(\mathbb{N}) = L^p(\mathbb{N},c) = \left\{ (a_1,a_2,\dots) : a_i \in \mathbb{C}, \sum_{i=1}^{\infty} |a_i|^p < \infty \right\},$$

(here we implicitly identify the function $f : \mathbb{N} \to \mathbb{K}$ with a sequence $(f(1),f(2),\dots)$) again a concrete object that we study in functional analysis.

Theorem 6.2.16 (Duality). *Suppose $1 \leq p < \infty$, μ is a σ-finite positive measure on (X,Σ), and Φ is a bounded linear functional on $L^p(\mu)$. Then there is a unique $g \in L^q(\mu)$ (where q is the exponential conjugate to p[(7)]) such that for every $f \in L^p(\mu)$,*

$$\Phi(f) = \int_X fg\, d\mu \quad \textit{and} \quad \|\Phi\| = \|g\|_q.$$

[(7)]If $p > 1$, q satisfies $\frac{1}{p} + \frac{1}{q} = 1$. If $p = 1$, define $q = \infty$.

Here $L^p(\mu)$ is either a collection of complex functions or a collection of extended real-valued functions.

Proof. We first assume $\mu(X) < \infty$. For $E \in \Sigma$, let $\lambda(E) = \Phi(\chi_E)$, then λ is countably additive by dominated convergence theorem, hence λ is a complex measure on Σ. Also since

$$|\lambda(E)| \le \|\Phi\| \|\chi_E\|_p = \|\Phi\| \mu(E)^{1/p},$$

$\mu(E) = 0$ implies $\lambda(E) = 0$, meaning that $\lambda \ll \mu$. By Radon-Nikodym theorem there is a unique integrable $g \in L^1(\mu)$ such that $d\lambda = g\, d\mu$. Hence

$$\Phi(f) = \int_X f g \, d\mu \tag{6.2.17}$$

holds when f is a simple function, and hence holds for any $f \in L^\infty(\mu)$ since every bounded measurable function is a uniform limit of a sequence of simple functions. We now show that RHS of (6.2.17) is indeed continuous on $L^p(\mu)$ by looking the following two cases.

Case 1. When $p = 1$, for any measurable E with $\mu(E) > 0$, let $f = \chi_E$ in (6.2.17),

$$\int_E g \, d\mu \le |\Phi(\chi_E)| \le \|\Phi\| \mu(E),$$

hence $g \le \|\Phi\|$ a.e. so that $\|g\|_\infty \le \|\Phi\|$. We conclude $|\int_X f g \, d\mu| \le \|g\|_\infty \|f\|_1$.

Case 2. When $1 < p < \infty$, let $E_n = \{x \in X : |g(x)| \le n\}$, then $g|_{E_n}$ is bounded. Let $\alpha(x)$ be a measurable function such that $|\alpha(x)| = 1$ for each x and $\alpha g = |g|$. Let $f = \chi_{E_n} |g|^{q-1} \alpha$ in (6.2.17), then

$$\int_{E_n} |g|^q \, d\mu = \Phi(f) \le \|\Phi\| \|f\|_p = \|\Phi\| \left(\int_{E_n} |g|^q \right)^{1/p},$$

this implies

$$\left(\int_{E_n} |g|^q \, d\mu \right)^{1/q} \le \|\Phi\|.$$

As E_n is ascending and $g \ne \infty$ a.e., hence $\|g\|_q \le \|\Phi\|$. By Hölder's inequality (see Problem 5.17) $|\int_X f g \, d\mu| \le \|f\|_p \|g\|_q$ for every $f \in L^p(\mu)$.

Now both sides of (6.2.17) define bounded linear functionals on $L^p(\mu)$ which agree on a dense subspace $L^\infty(\mu)$ of $L^p(\mu)$, hence $\Phi(f) = \int_X f g \, d\mu$ for every $f \in L^p(\mu)$. Moreover, $\|\Phi(f)\| \le \|f\|_p \|g\|_q$ implies $\|\Phi\| \le \|g\|_q$. Together with the bounds found in case 1 and 2, $\|\Phi\| = \|g\|_q$. We have completed the proof when X is a finite measure space.

Suppose $\mu(X) = \infty$, then since μ is σ-finite, by Lemma 6.1.17 there is a measurable w such that $0 < w < 1$ on X s.t. $d\tilde{\mu} := w \, d\mu$ defines a finite measure on X. Let $i : L^p(\tilde{\mu}) \to L^p(\mu)$ be defined by $i(f) = f w^{1/p}$, then i is an isometric isomorphism. Define $\Psi = \Phi \circ i$, then Ψ is a bounded linear functional on $L^p(\tilde{\mu})$, so by the preceding case we can find a $G \in L^q(\tilde{\mu})$ such that for every $F \in L^p(\tilde{\mu})$,

$$\Psi(F) = \int_X F G \, d\tilde{\mu}.$$

If $p = 1$, $\|\Phi\| = \|\Psi\| = \|G\|_{L^\infty(\tilde{\mu})} = \|G\|_{L^\infty(\mu)}$. Also if $p > 1$,

$$\|\Phi\| = \|\Psi\| = \left(\int_X |G|^q \, d\tilde{\mu} \right)^{1/q} = \left(\int_X |G w^{1/q}|^q \, d\mu \right)^{1/q}.$$

Hence we let $g = G$ if $p = 1$ and let $g = Gw^{1/q}$ if $p > 1$ such that $g \in L^q(\mu)$ and $\|\Phi\| = \|g\|_{L^q(\mu)}$. Finally for every $f \in L^p(\mu)$, $i^{-1}(f) = w^{-1/p} f$ and

$$\Phi(f) = \Psi(w^{-1/p} f) = \int_X f G w^{1-1/p}\, d\mu = \int_X f g\, d\mu. \qquad \square$$

By Theorem 6.2.16 when μ is σ-finite, positive and $1 \le p < \infty$ every bounded linear functional on $L^p(X,\mu)$ is of the form Φ_g defined by $\Phi_g(f) = \int_X f g\, d\mu$. Since $\|\Phi_g\| = \|g\|_q$, the function $i : L^q(\mu) \to (L^p(\mu))^*$ defined by $i(g) = \Phi_g$ is an isometric isomorphism, in that case we say that $(L^p(\mu))^* = L^q(\mu)$.

When $p = \infty$, generally we **don't** have $(L^\infty(\mu))^* = L^1(\mu)$. More specific, we will show that when $X = \mathbb{N}$ endowed with counting measure c, or when $X = \mathbb{R}$ endowed with Lebesgue measure m, the theorem fails (these are the important cases we usually use, at least for computation). Recall that

$$\ell^1 = \left\{ a = (a_1, a_2, \dots) : \|a\|_1 = \sum_{n=1}^{\infty} |a_n| < \infty \right\}$$

and

$$\ell^\infty = \{a = (a_1, a_2, \dots) : \|a\|_\infty = |\sup\{a_i : i = 1, 2, \dots\}| < \infty\}.$$

$\|\cdot\|_1$ and $\|\cdot\|_\infty$ are respectively the norms on them. We define $c_0 \subseteq \ell^\infty$ to be the sequences that converge to 0. Evidently $c_0 \subseteq \ell^\infty$. In what follows we assume the important extension result without proof: The Hahn-Banach theorem. The proof of this standard (not meaning easy!) result can be easily found in any related text.

Proposition 6.2.18. *Theorem 6.2.16 can fail when $p = \infty$ in the following two important cases:*

(i) ℓ^1 can be identified with a proper subspace of $(\ell^\infty)^*$.

(ii) $L^1(m)$ can be identified with a proper subspace of $(L^\infty(m))^*$.

Proof. (i) Define $T : \ell^1 \to (\ell^\infty)^*$ by $T(a)(b) = \sum_{n=1}^\infty a_n b_n$, which is an isometric embedding, as easily shown. Note that c_0 is a proper and closed subspace of ℓ^∞, by Hahn-Banach theorem there is a $f \in (\ell^\infty)^*$ such that $f|_{c_0} = 0$ and $f \neq 0$. We now show that T is not onto by showing that $T(a) \neq f$, for every $a \in \ell^1$. Assume $T(a) = f$ for some $a \in \ell^1$, then for any $b \in c_0$, $\sum_{n=1}^\infty a_n b_n = T(a)(b) = f(b) = 0$. Taking $b = e_i = (\underbrace{0, \dots, 0}_{i-1}, 1, 0, 0, \dots)$ for $i = 1, 2, \dots$, then $a_1 = a_2 = \cdots = 0$, showing that $f = 0$ on ℓ^∞, a contradiction.

(ii) Similarly consider the isometric embedding $i : L^1(m) \to (L^\infty(m))^*$ defined by $i(g)(f) = \int_{\mathbb{R}} f g\, dm$. To construct a functional on $L^\infty(m)$ that is not a image of i, we let

$$Y = \left\{ f \in L^\infty(m) : \lim_{r \to 0^+} \frac{1}{r} \int_{(0,r)} f\, dm \text{ exists} \right\}$$

which is a nonempty vector subspace of $L^\infty(m)$ and for $f \in Y$ we define $Lf = \lim_{r\to 0^+} \frac{1}{r} \int_{(0,r)} f\, dm$. Which is also a bounded linear functional on Y with $\|L\| = 1$. By Hahn-Banach theorem we can extend L to a functional $\tilde{L}$ on $L^\infty(m)$. If it happens that $i(g) = \tilde{L}$ for some $g \in L^1(m)$, then for any $x \in \mathbb{R}$, we can let $f = \chi_{(-\infty,x)}$ such that

$$\int_{(-\infty,x)} g\, dm = i(g)(\chi_{(-\infty,x)}) = \tilde{L}(\chi_{(-\infty,x)}) = L(\chi_{(-\infty,x)}) = \begin{cases} 0, & x \le 0, \\ 1, & x > 0. \end{cases}$$

Hence $g = 0$ m-a.e. (see Corollary 6.3.11), a contradiction since $L \neq 0$. □

6.3 Differentiation on Euclidean Space

6.3.1 Lebesgue Differentiation Theorem

In this subsection the ball $B(x,r) \subseteq \mathbb{R}^n$ is, as usual, induced by the 2-norm.

Definition 6.3.1. For every complex Borel measure λ on $\mathbb{R}^n$ we define the following quotient at x:

$$(Q_r\lambda)(x) = \frac{\lambda(B(x,r))}{m(B(x,r))},$$

where m is the Lebesgue measure on $\mathbb{R}^n$. We define the **symmetric derivative of λ at x** by

$$(D\lambda)(x) = \lim_{r\to 0}(Q_r\lambda)(x)$$

provided the limit exists. We shall study the function $D\lambda$ with the help of **maximal function** $M\lambda$. If λ is positive, define

$$(M\lambda)(x) = \sup_{r>0}(Q_r\lambda)(x).$$

If λ is complex Borel measure, we define $M\lambda = M|\lambda|$.

In general we can write $M\lambda$ and $M|\lambda|$ interchangeably, as they are, by definition, indeed the same. The convention here is for the sheer purpose that M acts as a function from L^1 to another space, we shall define the meaning of Mf later to elaborate this point.

An extended real-valued function f on $\mathbb{R}^n$ is said to be **lower semicontinuous** if for every $\alpha \in \mathbb{R}$, then set $\{f > \alpha\}$ is open. Where $\{f > \alpha\}$ is a simplified notation for $\{x \in \mathbb{R}^n : f(x) > \alpha\}$, we shall often make use of this kind of simplifications. We first show that the maximal functions are measurable as follows:

Proposition 6.3.2. *Let λ be a positive Borel measure on $\mathbb{R}^n$, the maximal function $M\lambda : \mathbb{R}^n \to [0,\infty]$ is lower semicontinuous.*

Proof. Let $x \in \{M\lambda > \alpha\}$, then $(M\lambda)(x) > \alpha$ implies there is $r > 0$ such that $\lambda(B(x,r)) = tm(B(x,r))$, for some $t > \alpha$. Let $\|y - x\| < \delta$, then $B(y, r+\delta) \supseteq B(x,r)$ and

$$\lambda(B(y,r+\delta)) \geq \lambda(B(x,r)) = tm(B(x,r)) = t\left(\frac{r}{r+\delta}\right)^n m(B(x,r+\delta)).$$

Since we can fix a $\delta > 0$ such that $t\left(\frac{r}{r+\delta}\right)^n > \alpha$, for this δ we conclude $\|y - x\| < \delta$ implies $(M\lambda)(y) \geq (Q_{r+\delta}\lambda)(y) > \alpha$, hence $\{M\lambda > \alpha\}$ is open. □

Our main interest is to the estimate given in Theorem 6.3.4, for this we need the following covering lemma.

Lemma 6.3.3. *Let C be a collection of open balls in $\mathbb{R}^n$ and let $U = \bigcup_{B\in C} B$. If $c < m(U)$, then there are disjoint $B_1, B_2, \ldots, B_n \in C$ such that $c < 3^n \sum_{i=1}^k m(B_i)$.*

Proof. Let $c < m(\bigcup_{V\in C} V)$, then there is a compact $K \subseteq \bigcup_{V\in C} V$ such that $c < m(K)$. There are $V_1,\dots,V_N \in C$ such that $K \subseteq \bigcup_{i=1}^N V_i$. Write $V_i = B(x_i,r_i)$. Relabel them if necessary, we assume $r_1 \ge r_2 \ge \cdots \ge r_n$.

Let $i_1 = 1$. For $i \ne 1$, discard those V_i's that have nonempty intersection with V_1 and let V_{i_2}, if any, be among the remaining ones that has largest radius. For $i \ne i_1,i_2$, discard those V_i's that have nonempty intersection with V_{i_2} and let V_{i_3}, if any, be among the remaining ones that has largest radius. The process terminates after finitely many steps and we get a disjoint collection $\{V_{i_1},V_{i_2},\dots,V_{i_k}\}$. A simple checking shows that

$$\bigcup_{i=1}^N V_i \subseteq \bigcup_{j=1}^k B(x_{i_j},3r_{i_j}) \implies c < m(K) \le 3^n \sum_{j=1}^k m(V_{i_j}). \qquad \square$$

Theorem 6.3.4. *If λ is a complex Borel measure on $\mathbb{R}^n$ and α is a positive number, then*

$$m\{M\lambda > \alpha\} \le 3^n \alpha^{-1}\|\lambda\|. \tag{6.3.5}$$

Proof. Let K be a compact subset of $\{M\lambda > \alpha\}$. For each $x \in K$, there is $r_x > 0$ such that $(Q_{r_x}\lambda)(x) = |\lambda|(B(x,r_x))/m(B(x,r_x)) > \alpha$. Now there are $x_1,x_2,\dots,x_n$ such that $K \subseteq \bigcup_{i=1}^n B(x_i,r_{x_i})$, hence

$$m(K) \le m\left(\bigcup_{i=1}^n B(x_i,r_{x_i})\right).$$

Let $t < m(K)$, by Lemma 6.3.3 there is a disjoint subcollection $\{B(x_{i_j},r_{x_{i_j}})\}_{j=1}^k$ such that $t < 3^n \sum_{j=1}^k m(B(x_i,r_{x_i})) \le 3^n\alpha^{-1}\sum_{j=1}^k |\lambda|(B(x_i,r_{x_i})) \le 3^n\alpha^{-1}\|\lambda\|$. Let $t \to m(K)^-$, we have

$$m(K) \le 3^n\alpha^{-1}\|\lambda\|.$$

Since this is true for each compact subset of $\{M\lambda > \alpha\}$, (6.3.5) follows. $\square$

Definition 6.3.6. We associate each $f \in L^1(\mathbb{R}^n,m)$ a **(Hardy-Littlewood) maximal function** define by

$$(Mf)(x) = (M(f\,dm))(x).$$

By (6.3.5) for every $\alpha > 0$,

$$m\{Mf > \alpha\} \le 3^n\alpha^{-1}\|f\,dm\| = 3^n\alpha^{-1}\|f\|_1.$$

The estimate roughly shows that if the total variation of f from zero relative to the whole space is small, then the place at which f varies from zero largely in a relative scale must also be small.

Definition 6.3.7. If $f \in L^1(\mathbb{R}^n)$, any $x \in \mathbb{R}^n$ for which it is true that

$$\lim_{r\to 0}\frac{1}{m(B(x,r))}\int_{B(x,r)} |f(y)-f(x)|\,dm(y) = 0$$

is called a **Lebesgue point of f**.

For example, it is obvious that every point of continuity of f is a Lebesgue point. It is not so obvious every $f \in L^1(\mathbb{R}^n)$ has Lebesgue point, the following result shows that they exist almost ubiquitously.

Theorem 6.3.8 (Lebesgue Differentiation). *If $f \in L^1(\mathbb{R}^n)$, then almost every $x \in \mathbb{R}^n$ is a Lebesgue point of f.*

Proof. Define

$$A_r f(x) = \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dm(y).$$

We also define $Af(x) = \overline{\lim}_{r\to 0} A_r f(x)$. Let $\epsilon > 0$ be given, since the collection of continuous functions on $\mathbb{R}^n$ is dense in $L^1(m)$ (see Problem 5.41), we can find a continuous g on $\mathbb{R}^n$ such that $\|f - g\|_1 < \epsilon$. Define $h = f - g$, then $\|h\|_1 < \epsilon$. Now

$$A_r f(x) \le A_r h(x) + A_r g(x) \le (Mh)(x) + |h(x)| + A_r(g)(x),$$

by taking $\overline{\lim}_{r\to 0}$ on both sides,

$$Af(x) \le Mh(x) + |h(x)|.$$

Now if we fix an $\alpha > 0$,

$$\{Af > 2\alpha\} \subseteq \{Mh > \alpha\} \cup \{|h| > \alpha\},$$

hence $m\{Af > 2\alpha\} \le (3^n + 1)\alpha^{-1}\|h\|_1 < (3^n + 1)\alpha^{-1}\epsilon$. Since $\epsilon > 0$ can be fixed arbitrarily, hence $m\{Af > 2\alpha\} = 0$, showing that $Af = 0$ m-a.e.. ☐

A more useful and general form of Theorem 6.3.8 is in terms of the following special family of sets shrinking to x.

Definition 6.3.9. A family of Borel sets $\{E_r\}_{r>0}$ in $\mathbb{R}^n$ is said to **shrink nicely to $x \in \mathbb{R}^n$** if

(i) $E_r \subseteq B(x,r)$.

(ii) There is $\alpha > 0$ independent of r such that $m(E_r) \ge \alpha m(B(x,r))$.

Note that we do not require $x \in E_r$ in the definition. Although α does not depend on r, it does depend on x and sometimes we write $\alpha = \alpha(x)$ for emphasis. For example, on $\mathbb{R}$ the collection $\{(x, x+r)\}_{r>0}$ shrinks nicely to x. On $\mathbb{R}^n$ let U be any Borel subset of $B(0,1)$ such that $m(U) > 0$, then the collection $\{x + rU\}_{r>0}$ also shrinks nicely to x.

Theorem 6.3.10 (Lebesgue Differentiation). *Let $f \in L^1(\mathbb{R}^n)$ and associate each x a family of sets $\{E_r(x)\}$ that shrinks nicely to x, then*

$$\lim_{r\to 0} \frac{1}{m(E_r(x))} \int_{E_r(x)} |f(y) - f(x)|\, dm(y) = 0$$

for every Lebesgue point x of f (hence m-a.e.).

Proof. For each Lebesgue point x, let $\alpha = \alpha(x)$ be defined as in Definition 6.3.9, then

$$\frac{1}{m(E_r(x))}\int_{E_r(x)} |f(y)-f(x)|\,dm(y) \le \frac{1}{\alpha(x)m(B(x,r))}\int_{B(x,r)} |f(y)-f(x)|\,dm(y),$$

the result follows from Theorem 6.3.8. □

In the study of Lebesgue measure on $\mathbb{R}$ there is usually an exercise stating that if $f \in L^1[a,b]$, for some $a,b \in \mathbb{R}$ and

$$\int_a^x f\,dm = 0$$

for every $x \in (a,b)$, then $f = 0$ m-a.e.. By the experience in a first course to mathematical analysis it is tempting to do differentiation (it is easy when f is continuous). When f is Riemann integrable, then $f = 0$ a.e. since f is continuous a.e.. More generally, Corollary 6.3.11 states that differentiation can also be done pointwise m-a.e. for Lebesgue integrable function.

Corollary 6.3.11. *Let $f \in L^1(a,b)$, where $a,b \in [-\infty,\infty], a \le b$. Define $F \in C[a,b]$ by*

$$F(x) = \int_{(a,x)} f\,dm,$$

then for m-a.e. x, $F'(x)$ exists and $F'(x) = f(x)$.

Proof. We may assume $f \in L^1(\mathbb{R})$ by setting $f|_{\mathbb{R}-(a,b)} = 0$, then $F(x) = \int_{(-\infty,x)} f\,dm$. Now

$$\frac{F(x+h)-F(x)}{h} = \begin{cases} \dfrac{1}{m([x,x+h])}\displaystyle\int_{[x,x+h]} f\,dm, & h \ge 0,\\ \dfrac{1}{m([x+h,x])}\displaystyle\int_{[x+h,x]} f\,dm, & h < 0.\end{cases}$$

Since both $\{[x,x+h]\}_{h>0}$ and $\{[x-h,x]\}_{h>0}$ shrink nicely to x, hence for every Lebesgue point x of f, $F'(x)$ exists and equals to $f(x)$. □

Now we return to the study of $D\lambda$. We now show that it is possible to compute the Radon-Nikodym derivative when $\lambda \ll m$ is a complex Borel measure on $\mathbb{R}^n$.

Theorem 6.3.12. *Let λ be a complex Borel measure on $\mathbb{R}^n$ and $\lambda \ll m$. Then $D\lambda = d\lambda/dm$ m-a.e..*

Proof. Radon-Nikodym asserts that $d\lambda/dm \in L^1(\mathbb{R}^n)$. At every Lebesgue point x of $d\lambda/dm$,

$$(D\lambda)(x) = \lim_{r\to 0}\frac{\lambda(B(x,r))}{m(B(x,r))} = \lim_{r\to 0}\frac{1}{m(B(x,r))}\int_{B(x,r)}\frac{d\lambda}{dm}\,dm = \frac{d\lambda}{dm}(x). \qquad \square$$

The differentiation of absolutely continuous measures is understood, next we deal with measures that are singular w.r.t. m. This can be done with the estimate given in Lemma 6.3.19. Before that we need to introduce:

Definition 6.3.13. Let λ be a positive Borel measure on $\mathbb{R}^n$. Define the **upper derivative of λ at x** by

$$(\overline{D}\lambda)(x) = \overline{\lim_{r\to 0}}(Q_r\lambda)(x).$$

Evidently $\overline{D}\lambda \le M\lambda$. $\overline{D}\lambda$ is also measurable. To see this, since

$$(\overline{D}\lambda)(x) = \lim_{n\to\infty}\Big(\sup_{0<r<1/n}(Q_r\lambda)(x)\Big),$$

each $\sup_{0<r<1/n}(Q_r\lambda)(x)$ is a measurable function in x by exactly the same argument as in Proposition 6.3.2, and thus $\overline{D}\lambda$ is measurable. To show $\overline{D}\lambda$ vanishes somewhere, we need the estimate given by Lemma 6.3.19. To prove this, we need an elementary result that justifies any finite positive Borel measure on $\mathbb{R}^n$ is regular. As usual the Borel σ-algebra on X is denoted by $\mathcal{B}_X$.

Lemma 6.3.14. *Let (X,d) be a complete separable metric space and let μ be a finite positive measure on $\mathcal{B}_X$, then μ is regular: For every $E\in\mathcal{B}_X$,*

$$\begin{aligned}\mu(E) &= \inf\{\mu(U) : U \supseteq E \text{ is open}\} && (6.3.15)\\ &= \sup\{\mu(L) : L\subseteq E \text{ is closed}\} && (6.3.16)\\ &= \sup\{\mu(K) : K\subseteq E \text{ is compact}\}. && (6.3.17)\end{aligned}$$

(6.3.15) and (6.3.16), as we shall see, do not require the completeness and separability of X (but finiteness of μ is crucial). The most useful one is (6.3.17) and conditions like completeness and separability will come into play. Note that for μ, both (6.3.15) and (6.3.16) hold if and only if

$$\begin{gathered}\text{for any given } \epsilon>0, \text{ there are a closed set } L \text{ and an open}\\ \text{set } U \text{ such that } L\subseteq E\subseteq U \text{ and } \mu(U-L)<\epsilon.\end{gathered} \tag{6.3.18}$$

We leave this easy verification as an exercise. To show (6.3.18) holds for each Borel set we follow the usual strategy: Our first step is to show that the collection of sets for which (6.3.18) holds forms a σ-algebra, and the second step is to show that all closed sets satisfy (6.3.18).

Proof. We show that each Borel set satisfies (6.3.18) first. Of course $\emptyset$ satisfies (6.3.18). Let $E_1, E_2, \ldots$ be subsets of X for which (6.3.18) holds, then $X-E_1$ satisfies (6.3.18) immediately. Let $\epsilon>0$ be given, then there are a closed set L_i and an open set U_i such that $L_i\subseteq E_i\subseteq U_i$ and $\mu(U_i-L_i)<\epsilon/2^{i+1}$. Since μ is finite,

$$\lim_{n\to\infty}\mu\left(\bigcup_{i=1}^{\infty}E_i-\bigcup_{i=1}^{n}L_i\right)=\mu\left(\bigcup_{i=1}^{\infty}E_i-\bigcup_{i=1}^{\infty}L_i\right)\le\sum_{i=1}^{\infty}\mu(E_i-L_i)<\frac{\epsilon}{2}.$$

Hence we can find an n such that $\mu(\bigcup E_i-\bigcup_{i=1}^n L_i)<\epsilon/2$. Let $U=\bigcup_{i=1}^{\infty}U_i$ and $L=\bigcup_{i=1}^{n}L_i$,

$$\mu(U-L)\le\mu(U-\bigcup E_i)+\mu(\bigcup E_i-L)<\frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon.$$

Hence $\bigcup E_i$ also satisfies (6.3.18). We conclude those sets satisfying (6.3.18) forms a σ-algebra. Since each closed set L in X is a countable intersection of open sets, closed sets satisfy (6.3.18), and our two steps are completed.

To show (6.3.17), let $\epsilon > 0$ be given and L a closed subset of $E \in \mathcal{B}_X$ such that $\mu(E-L) < \epsilon$. Let $C := \{c_1, c_2, \dots\} \subseteq X$ be a countable dense subset. For each $\alpha > 0$, $X = \bigcup B(c_i, \alpha)$ as C is dense. Let $\alpha = 1/n$, where $n \in \mathbb{N}$, write $\overline{B}(x,r) = \{y \in X : d(y,x) \le r\}$ and write

$$L = \bigcup(\overline{B}(c_i, 1/n) \cap L).$$

We can find a $k_n \in \mathbb{N}$ such that

$$\mu\left(L - \bigcup_{i=1}^{k_n}(\overline{B}(c_i, 1/n) \cap L)\right) < \frac{\epsilon}{2^n}.$$

These closed subsets are already very close inner approximation of L, so is $K := \bigcap_{n=1}^{\infty} \bigcup_{i=1}^{k_n}(\overline{B}(c_i, 1/n) \cap L)$, which is closed and totally bounded and hence compact. Moreover,

$$\mu(L-K) \le \sum_{n=1}^{\infty} \mu\left(L - \bigcup_{i=1}^{k_n}(\overline{B}(c_i, 1/n) \cap L)\right) < \epsilon,$$

hence $\mu(E-K) \le \mu(E-L) + \mu(L-K) < \epsilon + \epsilon = 2\epsilon$, thus (6.3.17) holds. □

Lemma 6.3.19. *Let λ be a finite positive Borel measure on $\mathbb{R}^n$, $E \in \mathcal{B}_{\mathbb{R}^n}$ and $\alpha > 0$, then*

$$m\{x \in E : (\overline{D}\lambda)(x) > \alpha\} \le 3^n \alpha^{-1} \lambda(E). \tag{6.3.20}$$

Proof. We imitate the proof of Theorem 6.3.4 by considering a compact set K and an open set V such that $K \subseteq \{x \in E : (\overline{D}\lambda)(x) > \alpha\}$ and $V \supseteq E$. We further assume those balls in the proof satisfy $B(x, r_x) \subseteq V$, now we repeat the proof and arrive to

$$m(K) \le 3^n \alpha^{-1} \lambda(V).$$

Since $\mathbb{R}^n$ is a complete separable metric space and λ is finite and positive on $\mathcal{B}_{\mathbb{R}^n}$, by Lemma 6.3.14 λ is regular, (6.3.20) follows. □

Corollary 6.3.21. *Associate to each $x \in \mathbb{R}^n$ a family $\{E_r(x)\}_{r>0}$ that shrinks to x nicely. If λ is a complex Borel measure and $\lambda \perp m$, then for m-a.e. $x \in \mathbb{R}^n$,*

$$\lim_{r \to 0} \frac{\lambda(E_r(x))}{m(E_r(x))} = 0. \tag{6.3.22}$$

Proof. Since we can express λ as a linear span of two finite signed measures, and each finite signed measure have a Hahn decomposition. So it is enough to prove the case that λ is finite positive Borel measure on $\mathbb{R}^n$, let's assume that is the case. Since

$$\frac{\lambda(E_r(x))}{m(E_r(x))} \le \frac{\lambda(B(x,r))}{\alpha(x) m(B(x,r))},$$

it is also enough to prove the case that $E_r(x) = B(x,r)$.

As $\lambda \perp m$, there are disjoint Borel sets A, B such that $A \sqcup B = \mathbb{R}^n$, $\lambda \to A$ and $m \to B$. Since $\lambda(B) = 0$, Lemma 6.3.19 shows that $D\lambda = 0$ m-a.e. on B. Since $m(A) = 0$, (6.3.22) holds m-a.e.. □

We combine the results so far to conclude:

Theorem 6.3.23. *Associate to each $x \in \mathbb{R}^n$ a family $\{E_r(x)\}_{r>0}$ that shrinks to x nicely. Let λ be a complex Borel measure and $d\lambda = f\,dm + d\lambda_s$ the Lebesgue decomposition of λ w.r.t. m, then for m-a.e. $x \in \mathbb{R}^n$,*

$$\lim_{r\to 0}\frac{\lambda(E_r(x))}{m(E_r(x))} = f(x) \quad \text{and} \quad D\lambda_s(x) = 0.$$

6.3.2 Application

We can construct a continuous function on an open interval which fails to be differentiable at any point (due to Karl Weierstrass), so we would ask: Which kind of functions can be differentiable, say, at least one point? With the machineries developed so far this question can be quickly answered!

Theorem 6.3.24. *Let $f : \mathbb{R} \to \mathbb{R}$ be increasing and let $F(x) = f(x^+)$. Then $F(x)$ is differentiable m-a.e. and if $F'(x)$ exists, so does $f'(x)$ and $f'(x) = F'(x)$.*

Proof. Since F is right continuous, we associate F with a Lebesgue-Stieltjes measure μ_F. Let $k \in \mathbb{N}$, then $\mu_F'(E) := \mu_F(E \cap (k,k+1))$ is a finite Borel measure on $\mathbb{R}$. Let $h_n > 0$, $\lim_{n\to\infty} h_n = 0$ and $x \in (k,k+1)$, then for large enough n,

$$\frac{F(x+h_n)-F(x)}{h_n} = \frac{\mu_F'\big((x,x+h_n)\big)}{m\big((x,x+h_n)\big)}. \tag{6.3.25}$$

Theorem 6.3.23 asserts that RHS of (6.3.25) exists for m-a.e. $x \in \mathbb{R}$, so $\lim_{n\to\infty}\frac{F(x+h_n)-F(x)}{h_n}$ exists m-a.e. on $(k,k+1)$. Likewise for $k_n < 0$ and $\lim_{n\to\infty} k_n = 0$, the limit $\frac{F(x+k_n)-F(x)}{k_n}$ exists m-a.e. on $(k,k+1)$, we conclude $F'(x)$ exists m-a.e. on $(k,k+1)$. Since this is true for each fixed $k \in \mathbb{Z}$, $F'(x)$ exists m-a.e. on $\mathbb{R}$.

Next we give an elementary proof to second half of the theorem which is due to me: Assume $F'(x)$ exists for some $x \in (a,b)$.

Claim. *f is continuous at x.*

Proof. Let $\epsilon,\delta > 0$ be given, by continuity of F we can find $x' < x$ such that $|x'-x| < \delta$ and $|F(x')-F(x)| < \epsilon/2$. Next, by right-limit definition we can find x'' close to x' such that $x' < x'' < x$ and $|F(x')-f(x'')| < \epsilon/2$, and also $x-x'' < x-x' < \delta$. Combining these two estimates, we have shown that given $\epsilon,\delta > 0$, there is $x'' < x$ with $|x''-x| < \delta$ such that

$$|F(x)-f(x'')| \le |F(x)-F(x')| + |F(x')-f(x'')| < \epsilon.$$

Hence we choose $\epsilon = \delta = \frac{1}{n}$ and some $x'' = x_n < x$ so that $|x-x_n| < \frac{1}{n}$ and $|F(x)-f(x_n)| < \frac{1}{n}$, i.e., $f(x^-) = \lim_{n\to\infty} f(x_n) = F(x) = f(x^+)$, so f is continuous at x and $f(x) = F(x)$. ☐

Our focus now turns to existence of $f'(x)$. By the existence of $a := F'(x) = \lim_{y\to 0}\frac{F(x+y)-f(x)}{y}$, given $\epsilon > 0$ we can choose $\delta > 0$ (small enough that $(x-\delta,x+\delta) \subseteq [a,b]$) such that

$$\forall y \in (-\delta,\delta)-\{0\}, \quad \lim_{\varepsilon\to 0^+}\left|\frac{f(x+y+\varepsilon)-f(x)}{y+\varepsilon} - a\right| = \left|\frac{F(x+y)-f(x)}{y} - a\right| < \epsilon. \tag{6.3.26}$$

By the right-hand limit in (6.3.26),

$$\forall y \in (-\delta,\delta)-\{0\}, \exists \delta_y > 0, \forall \varepsilon \in (0,\delta_y), \quad \left|\frac{f(x+y+\varepsilon)-f(x)}{y+\varepsilon} - a\right| < \epsilon. \tag{6.3.27}$$

Here we choose δ_y small such that $\delta_y < \min\{|y|, \delta - y\}$ to avoid $y+\varepsilon = 0$ and ensure $y + \delta_y < \delta$ (we want everything happens in $(-\delta,\delta)$). Define

$$A = \left\{ z \in [a,b] : \left| \frac{f(x+z)-f(x)}{z} - a \right| < \epsilon \right\},$$

by (6.3.27) for each $y \in (-\delta,\delta)-\{0\}$, $y + (0,\delta_y) = (y, y+\delta_y) \subseteq A$, hence

$$U := \bigcup_{y \in (-\delta,\delta)-\{0\}} (y, y+\delta_y) \subseteq A.$$

Now $U \subseteq (-\delta,\delta)$ is open, there are $a_i, b_i \in (-\delta,\delta)$ such that $U = \bigsqcup (a_i, b_i)$, where i's are positive integers.

Claim. *Let $E = \{0, a_i : \text{all } i\}$, we have $(-\delta,\delta) - E \subseteq U$.*

Proof. Let $x \in (-\delta,\delta) - E$, then $(x, x+\delta_x) \subseteq U$ implies $(x, x+\delta_x) \subseteq (a_i, b_i)$, for some i. As $x \notin E$, $x \neq a_i$, we have $a_i < x$, so $x \in U$. □

Claim. *$A' := A \cap (-\delta,\delta)$ is dense in $(-\delta,\delta)$.*

Proof. As a result of the last claim, by $U \subseteq A'$ we have $(-\delta,\delta) - E \subseteq U \subseteq A'$, hence $E \supseteq (-\delta,\delta) - A'$, so undesired points are those in E. Since E is countable, $m(E) = 0$, this implies $m\big((-\delta,\delta) - A'\big) = 0$, so A' is dense in $(-\delta,\delta)$. □

Finally let $\delta_n \to 0$, consider

$$a_n := \frac{f(x+\delta_n)-f(x)}{\delta_n}.$$

We show a_n converges by considering two cases **(i)** all $\delta_n > 0$ and **(ii)** all $\delta_n < 0$, after that by combing these two cases we are done.

Let $c \in (0,1)$ be given. Assume the case (i), i.e., $\delta_n > 0$ for all n. Since $\delta_n \to 0$, there is an N such that $n > N \implies |\delta_n| < \delta$. Then by density of A' in $(-\delta,\delta)$, for each fixed $n > N$ we can find $\delta_n', \delta_n'' \in A$ such that $0 < \delta_n' < \delta_n < \delta_n''$ with

$$\delta_n - \delta_n' < c\delta_n \iff (1-c)\delta_n < \delta_n'$$

and

$$\delta_n'' - \delta_n < c\delta_n \iff \delta_n'' < (1+c)\delta_n.$$

Then $n > N$ implies

$$a_n \geq \frac{\delta_n'}{\delta_n} \frac{f(x+\delta_n')-f(x)}{\delta_n'} \geq (1-c)\frac{f(x+\delta_n')-f(x)}{\delta_n'} \geq (1-c)(a-\epsilon)$$

and

$$a_n \leq \frac{\delta_n''}{\delta_n} \frac{f(x+\delta_n'')-f(x)}{\delta_n''} \leq (1+c)\frac{f(x+\delta_n'')-f(x)}{\delta_n''} \leq (1+c)(a+\epsilon).$$

Combining these two, we conclude

$$n > N \implies (1-c)(a-\epsilon) \le a_n \le (1+c)(a+\epsilon). \tag{6.3.28}$$

As $c \in (0,1)$ can be fixed arbitrarily, hence (6.3.28) becomes

$$n > N \implies a-\epsilon \le a_n \le a+\epsilon,$$

we conclude $\lim_{n\to\infty} a_n = a$. The case (ii) that $\delta_n < 0$ is essentially the same. $\square$

Chapter 7

Locally Compact Hausdorff Spaces and Riesz Representation Theorem

In this chapter given a topological space, "open" and "closed" are always with respect to the topology of the *largest space* in our discussion, unless otherwise specified. Also a first course in point-set topology is assumed.

7.1 Normal Spaces

7.1.1 Separation Properties

Definition 7.1.1. We say that U is a **neighborhood of a point x** if U is open and $x \in U$. We say that U is a **neighborhood of a subset K** if U is open and $U \supseteq K$. Moreover, we have the following separation properties:

(i) **Tychonoff** For every pair of distinct points $u_1, u_2 \in X$, there are neighborhood U_i of u_i such that $u_1 \notin U_2$ and $u_2 \notin U_1$.

(ii) **Hausdorff** Every two points can be separated by disjoint neighborhoods.

(iii) **Regular** The Tychonoff properties holds. Moreover, each closed set K and each point $x \notin K$ can be separated by disjoint neighborhoods.

(iv) **Normal** The Tychonoff properties holds. Moreover, each pair of disjoint closed sets can be separated by disjoint neighborhoods.

These are "adjective" of topological spaces having the respective separation properties. For example, each metric space is a Hausdorff topological space. Moreover, after Proposition 7.1.2 we can see that:

$$\mathcal{T}_{\text{normal}} \subseteq \mathcal{T}_{\text{regular}} \subseteq \mathcal{T}_{\text{Hausdorff}} \subseteq \mathcal{T}_{\text{Tychonoff}}.$$

Proposition 7.1.2. *A topological space X is Tychonoff iff every singleton in X is closed.*

Proof. $\{x\}$ is closed for all $x \in X$ iff $X - \{x\}$ is open for all $x \in X$ iff for all $x \in X$ and each $y \in X - \{x\}$, there is open set U such that $y \in U$ but $x \notin U$. □

Definition 7.1.3. In a topological space X we say that a set K has **Nested Neighborhood Property (NNP)** if for every open set $U \supseteq K$, there is open set O s.t. $K \subseteq O \subseteq \overline{O} \subseteq U$.

Proposition 7.1.4. *Let X be a Tychonoff topological space. Then X is normal iff every closed set in X has NNP.*

Proof. Assume X is normal. Let K be closed in X and $U \supseteq K$ be open. Then $X - U$ and K are disjoint closed subsets of X, by normality there are disjoint open sets V', V such that $V' \supseteq X - U$ and $V \supseteq K$. Since $V' \cap V = \emptyset$, one has $V \subseteq X - V'$, so

$$K \subseteq V \subseteq \overline{V} \subseteq X - V' \subseteq U,$$

that means K has NNP. Conversely, assume each closed set has NNP. Let H, K be disjoint closed subsets of X, then $H \subseteq X - K$, so there is open O, $H \subseteq O \subseteq \overline{O} \subseteq X - K$. Clearly $U := O \supseteq H$ and $V := X - \overline{O} \supseteq K$ and $U \cap V = \emptyset$, so the normal separation property holds. □

7.1.2 Compact Topological Spaces

We recall the definition of compactness first.

Definition 7.1.5. A topological space X is **compact** if any open cover of X have a finite subcover. A subset K of a topological space Y is **compact** if K is compact with respect to the subspace topology induced by Y.

The compactness of a subset K of Y can be rephrased as follows: Any open cover of K in Y has a finite subcover.

Proposition 7.1.6. *Let X be a topological space and K compact. Then any closed subset of K is again compact.*

Proof. Let $L \subseteq K$ and $\{U_\alpha\}$ be an open cover of L. Then $\bigcup U_\alpha \supseteq L = K \cap L$ and $X - L \supseteq K - L$, so $\bigcup U_\alpha \cup (X - L) \supseteq K$. Since $\{U_\alpha\} \cup \{X - L\}$ is an open cover of K, there is α_i such that $\bigcup_{i=1}^n U_{\alpha_i} \cup (X - L) \supseteq K \supseteq L$. As $(X - L) \cap L = \emptyset$, $\bigcup_{i=1}^n U_{\alpha_i} \supseteq L$. □

Proposition 7.1.7. *A compact subset K of a Hausdorff topological space X is closed in X.*

Proof. We appeal to some "finite property" of compactness, i.e., we try to make an open cover of K, by any means. To show K is closed, it amounts to show $X - K$ is open. Fix $y \in X - K$, for each $x \in K$, there are open sets $U_x \ni x$, $V_x \ni y$ such that $U_x \cap V_x = \emptyset$. Since $\{U_x\}_{x \in K}$ covers K, there are $x_1, \ldots, x_n \in K$ such that $U := \bigcup_{i=1}^n U_{x_i} \supseteq K$. Define $V = \bigcap_{i=1}^n V_{x_i}$, then $V \cap U = \emptyset$, meaning that $V \subseteq X - U \subseteq X - K$. Since the choice of y can be relaxed, $X - K$ is open. □

Proposition 7.1.8. *A compact Hausdorff space is normal.*

Proof. A Hausdorff space X is already Tychonoff. By Proposition 7.1.4, it suffices to show each closed subset has NNP. Let K be closed subset and let $U \supseteq K$ be open. Fix a $y \in X - U$, then $y \in X - K$, by the proof of Proposition 7.1.7 there are open sets $U_y \supseteq K$ and $V_y \ni y$ so that $U_y \cap V_y = \emptyset$. As $X - U$ is closed subset of a compact space X, $X - U$ is compact. Since $\{V_y\}_{y \in X-U}$ covers $X - U$, there are $y_1, \dots, y_n \in X - U$ so that $V := \bigcup_{i=1}^n V_{y_i} \supseteq X - U$. Define $O = \bigcap_{i=1}^n U_{y_i}$, then $O \cap V = \emptyset$, so that $O \supseteq K$ and $\overline{O} \subseteq \overline{X - V} = X - V \subseteq U$. $\square$

We will study locally compact Hausdorff (LCH) spaces and the measure theory on such spaces. As suggested by the name, an LCH space is locally a compact Hausdorff space, i.e., a normal space. So we can investigate LCH spaces by the tools on normal space, which we introduce in the next section.

7.1.3 Fundamental Theorems on Normal Spaces: Urysohn's Lemma and Tietze Extension Theorem

Urysohn's Lemma

Recall that a space is normal iff it is Tychonoff and every pair of disjoint closed sets can be separated by disjoint neighborhoods.

In a metric space M, for any pair of disjoint closed subsets A and B there is a function $f : X \to \mathbb{R}$ that takes value 0 on A and 1 on B, one such choice is

$$f_{A,B}(x) = \frac{d(x,B)}{d(x,A) + d(x,B)}. \tag{7.1.9}$$

The function is properly defined since $d(x,A) + d(x,B) = 0$ iff $x \in A \cap B$, which never happens as A, B are disjoint. This is a kind of extension result if we view in the following way: Let A be a closed subset of a metric space M, let $U \supseteq A$ be open, then $A, M - U$ are disjoint closed subsets of M, hence any constant function on A can be continuously extended to X, with the help of $f_{X-U,A}$.

We will prove the following result that extends our discussion.

Lemma 7.1.10 (Urysohn). *Let X be a normal topological space and A, B be (nonempty) disjoint closed subsets of X. For any closed interval $[a,b]$, there is a continuous function $f : X \to [a,b]$ such that $f|_A \equiv a$ and $f|_B \equiv b$.*

We have seen in Proposition 7.1.8 that compact Hausdorff spaces are normal. There are much more normal spaces:

Example 7.1.11. Metric spaces are normal. To see this, let A, B be pair of disjoint closed subsets of a metric space M and consider the function defined in equation (7.1.9), one has $f_{A,B}^{-1}(1) = A$ and $f_{A,B}^{-1}(0) = B$, so $U := f^{-1}(\frac{1}{2}, \frac{3}{2})$ and $V := f^{-1}(-\frac{1}{2}, \frac{1}{2})$ are disjoint neighborhoods of A and B. $\square$

We need some terminology to begin with.

Definition 7.1.12. Let X be a topological space and let $\Lambda \subseteq \mathbb{R}$. A collection of open sets $\{O_\lambda\}_{\lambda \in \Lambda}$ is said to be **normally ascending** if for every $\lambda_1, \lambda_2 \in \Lambda$,

$$\lambda_1 < \lambda_2 \implies \overline{O_{\lambda_1}} \subseteq O_{\lambda_2}.$$

Such collection defined in Definition 7.1.12 appears when we apply Proposition 7.1.4 several times. To see this, let's begin to prove Urysohn's lemma (recall the setting!).

Proof of Urysohn's Lemma 7.1.10. Let A,B be disjoint closed subsets of a normal space X, then $A \subseteq X - B =: U$. By normality, A has NNP and we index those open subsets of U (from NNP) in a special way as follows:

$$A \subseteq O_{1/2} \subseteq \overline{O_{1/2}} \subseteq U; \tag{7.1.13}$$

$$A \subseteq O_{1/4} \subseteq \overline{O_{1/4}} \subseteq O_{1/2} \subseteq \overline{O_{1/2}} \subseteq O_{3/4} \subseteq \overline{O_{3/4}} \subseteq U; \tag{7.1.14}$$

$$\vdots$$

In this way we have inductively defined $\{O_\lambda : \lambda \in \Lambda\}$, where

$$\Lambda = \left\{ \frac{m}{2^k} : m = 1,2,\dots,2^k - 1, k \geq 1 \right\}. \tag{7.1.15}$$

Clearly $\{O_\lambda\}_{\lambda\in\Lambda}$ defined here is normally ascending. Λ is called dyadic rationals in $(0,1)$, which is dense in $(0,1)$ since for each $x \in (0,1)$ we can define $\{[2^n x]/2^n\}_{2^n \geq 1/x}$.

It suffices to prove the case that $a = 0$ and $b = 1$. After we have obtained such f, then $g := (b-a)f + a$ will be the desired function in the lemma. We now construct such f.

We continue what we have done from (7.1.13) to (7.1.15). For each $x \in (0,1)$, define

$$f(x) = \begin{cases} \inf\{\lambda \in \Lambda : x \in O_\lambda\}, & x \in \bigcup_{\lambda\in\Lambda} O_\lambda, \\ 1, & x \in X - \bigcup_{\lambda\in\Lambda} O_\lambda. \end{cases}$$

It is clear that $f|_A \equiv 0$. Moreover, since $B \subseteq X - \bigcup_{\lambda\in\Lambda} O_\lambda$, $f|_B \equiv 1$. It suffices to check that f is continuous, which is true by Lemma 7.1.16. □

Lemma 7.1.16. *Let X be a topological space and Λ be a dense subset of (a,b). Let the collection of open subsets of X, $\{O_\lambda\}_{\lambda\in\Lambda}$, be normally ascending. Define $f : X \to [a,b]$ as follows:*

$$f(x) = \begin{cases} \inf\{\lambda \in \Lambda : x \in O_\lambda\}, & x \in \bigcup_{\lambda\in\Lambda} O_\lambda, \\ b, & x \in X - \bigcup_{\lambda\in\Lambda} O_\lambda. \end{cases}$$

Then f is continuous.

Proof. It is enough to prove that for every $c \in (a,b)$,

$$U := \{x \in X : f(x) < c\} \quad \text{and} \quad V := \{x \in X : f(x) > c\}$$

are open, after that $f^{-1}(O)$ is open for every set O open in $[a,b]$.

Let $x \in (a,b)$, then $f(x) < c$ if and only if $x \in \bigcup_{\lambda\in\Lambda} O_\lambda$ and there is $\lambda \in \Lambda$ such that $x \in O_\lambda$ and $\lambda < c$, which is the same as saying

$$x \in \left(\bigcup_{\lambda\in\Lambda} O_\lambda \right) \cap \left(\bigcup_{\lambda\in\Lambda\cap(a,c)} O_\lambda \right),$$

hence U is open. Next to show V is open, we observe that

$$f(x) = \begin{cases} \sup\{\lambda \in \Lambda : x \notin O_\lambda\}, & x \in X - \bigcap_{\lambda\in\Lambda} O_\lambda, \\ a, & x \in \bigcap_{\lambda\in\Lambda} O_\lambda. \end{cases}$$

We leave this as a simple exercise (to prove the equality, density of Λ is needed). Therefore we can give a similar reasoning as in showing U is open to get

$$V = \left(X - \bigcap_{\lambda\in\Lambda} O_\lambda\right) \cup \left(\bigcup_{\lambda\in\Lambda\cap(c,b)} O_\Lambda\right).$$

Due to normally ascending property and density of Λ, we have

$$\bigcap_{\lambda\in\Lambda} O_\lambda = \bigcap_{\lambda\in\Lambda} \overline{O_\lambda},$$

which is closed, hence V is open and thus we are done. $\square$

Tietze Extension Theorem

Now we apply Urysohn's lemma to prove a strong extension result.

Theorem 7.1.17 (Tietze Extension). *Let X be a normal topological space, K a closed subset of X and $f : K \to \mathbb{R}$ a continuous function.*

(i) f has a continuous extension $F : X \to \mathbb{R}$.

(ii) If f is bounded such that $f(X) \subseteq [a,b]$, it is possible to choose F in (i) so that $F(X) \subseteq [a,b]$.

It suffices to show (ii), (i) will then follow from (ii). This is because when f is unbounded, we can choose a homeomorphism $h : \mathbb{R} \to (-1,1)$ so that $h \circ f$ is bounded, we may apply (ii) to extend $h \circ f$ to F, and $h^{-1} \circ F$ will extend f.

Proof. We first assume f is bounded. We may assume $a = \inf f(K)$ and $b = \sup f(K)$. We also assume $a = -1$ and $b = 1$[(1)].

Now assume that $f : K \to [-1,1]$ is continuous so that $-1 = \inf f(K)$ and $1 = \sup f(K)$. We first need a general fact:

Claim. *Let $h : K \to \mathbb{R}$ be continuous so that $|h| \le c$ where $-c = \inf h(K)$, $c = \sup h(K)$, then there is $g : X \to [-c,c]$ so that*

$$|g| \le \frac{c}{3} \text{ on } X \quad \text{and} \quad |h-g| \le \frac{2c}{3} \text{ on } K, \tag{7.1.18}$$

with $\inf_{x\in K}(h-g)(x) = -\frac{2c}{3}$ and $\sup_{x\in K}(h-g)(x) = \frac{2c}{3}$.

Proof. Consider $A := h^{-1}[-c,-\frac{c}{3}]$ and $B := h^{-1}[\frac{c}{3},c]$. Since $-c = \inf h(K)$, A is nonempty, so is B due to similar reason. Moreover, A,B are closed in K, but K is closed in X, so A,B are closed in X. Now by Urysohn's lemma there is a $g : X \to [-\frac{c}{3},\frac{c}{3}]$, $g|_A \equiv -\frac{c}{3}$ and $g|_B \equiv \frac{c}{3}$. It is easy to check (7.1.18) is satisfied. The infimum and supremum assertions can be proved by considering $h-g$ on A and B respectively. $\square$

[(1)]Otherwise we choose a canonical homeomorphism $h : [a,b] \to [-1,1]$ by $h(x) = \frac{2}{b-a}x - \frac{a+b}{b-a}$ and let $\tilde{f} := h \circ f : K \to [-1,1]$, extending $\tilde{f}$ in desired way is the same as extending f in desired way, since if $\tilde{f}$ extends to $\tilde{F}$ with $|\tilde{F}| \le 1$, then $-1 \le \tilde{F} \le 1 \implies a = h^{-1}(-1) \le h^{-1} \circ F \le h^{-1}(1) = b$, which extends f.

Now we let $h = f$ and $c = 1$ in our claim, then there is continuous g_1 on X such that $|g_1| \le \frac{1}{3}$ on X and $|f - g_1| \le \frac{2}{3}$ on K. Next apply $h = f - g_1$ and $c = \frac{2}{3}$, there is continuous g_2 on X such that $|g_2| \le \frac{1}{3}(\frac{2}{3})$ on X and $|f - g_1 - g_2| \le (\frac{2}{3})^2$ on K. Inductively, for each $n = 1, 2, \dots$ we can find g_n so that

$$|g_n| \le \frac{2^{n-1}}{3^n} \text{ on } X \quad \text{ and } \quad |f - g_1 - \cdots - g_n| \le \frac{2^n}{3^n} \text{ on } K.$$

By the first inequality, $s_n := \sum_{i=1}^n g_i$ converges uniformly to $g := \sum_{i=1}^\infty g_i$ on X. As each s_n is continuous, g is also continuous on X. By the second inequality, $f = g$ on K. Finally, $|g| \le \frac{1}{3} \sum_{i=1}^\infty \left(\frac{2}{3}\right)^{i-1} = \frac{1}{3} \cdot \frac{1}{1-\frac{2}{3}} = 1$. ☐

Remark. By requiring X in Theorem 7.1.17 be compact metric space, it is possible to prove Tietze Extension theorem by functional analysis approach. The idea is to show that for closed subspace $Y \subseteq X$ the linear map

$$F : C(X) \to C(Y); \quad f \mapsto f|_Y$$

is surjective which can be proved by tools developed on Banach spaces.

7.2 Locally Compact Hausdorff Spaces

We now study locally compact Hausdorff spaces by machineries developed on normal spaces.

Definition 7.2.1. A topological space X is said to be **locally compact** (or **LCH**) if each point in X has a neighborhood that has compact closure in X.

Definition 7.2.2. A subset A of a topological space X is said to be **precompact** if $\overline{A}$ is compact in X.

By using Definition 7.2.2, we can rephrase Definition 7.2.1 as: X is locally compact if each $x \in X$ has a precompact neighborhood.

Lemma 7.2.3. *Let X be an LCH space. If O is a neighborhood $x \in X$, then there is a precompact neighborhood V of x such that*

$$x \in V \subseteq \overline{V} \subseteq O.$$

Proof. Let $x \in O$, then x has a precompact neighborhood U. Since $\overline{U}$ is compact, hence normal. As $x \in O \cap U \subseteq \overline{U}$, by normality of $\overline{U}$ there is V open in $\overline{U}$ such that $x \in V \subseteq \overline{V}^{\overline{U}} \subseteq O \cap U$[(2)]. Since O and U are open, V is open in X and $\overline{V} \subseteq O$ is compact.☐

Proposition 7.2.4. *Let K be compact subset of an LCH space X and O a neighborhood of K, then there is a precompact neighborhood V of K such that*

$$K \subseteq V \subseteq \overline{V} \subseteq O.$$

[(2)]For $A \subseteq Y \subseteq X$, we denote $\overline{A}^Y$ the closure of A with respect to subspace topology of Y.

Proof. Let $k \in K$, then $k \in O$, by Lemma 7.2.3 there is a precompact neighborhood V_k of k such that
$$\{k\} \subseteq V_k \subseteq \overline{V_k} \subseteq O.$$
Since $K \subseteq \bigcup_{k\in K} V_k$, there is $k_1,\ldots,k_n \in K$ such that $K \subseteq \bigcup_{k=1}^n V_{k_i} \subseteq \bigcup_{k=1}^n \overline{V_{k_i}} \subseteq O$. Define $V = \bigcup_{k=1}^n V_{k_i}$, then V is open and precompact. □

For a real-valued function f on a topological space, we define the **support of f** by $\operatorname{spt} f = \overline{\{x \in X : f(x) \neq 0\}}$. We denote $C_c(X) = \{f \in C(X) : \operatorname{spt} f \text{ is compact}\}$ the collection of **compactly supported** continuous functions on X. A function $f \in C_c(X)$ if and only if it is continuous and vanishes outside a compact set. Since
$$\operatorname{spt}(f+g) \subseteq \operatorname{spt} f \cup \operatorname{spt} g,$$
$C_c(X)$ forms a vector space over $\mathbb{R}$. Note that we also have $\operatorname{spt}(fg) \subseteq \operatorname{spt} f \cap \operatorname{spt} g$ since $fg(x) \neq 0$ implies $f(x), g(x) \neq 0$ and $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$.

For example, the unshaded region in Figure 7.1 is $\operatorname{spt} f$ but f is not continuous. By the definition of closure, $\operatorname{spt} f$ is the smallest closed set outside of which f vanishes, and whenever $f(x) \neq 0$, $x \in \operatorname{spt} f$.

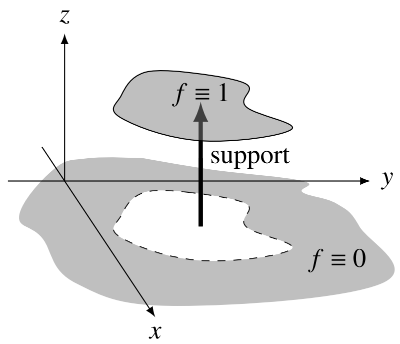


Figure 7.1: Support of f.

The following version of Urysohn's lemma tells us constant functions on LCH spaces as in the above picture can always be "smoothed"!

Lemma 7.2.5 (Urysohn, LCH Version). *Let X be an LCH space. Let K be compact and O a neighborhood of K, then there is $f \in C_c(X,[0,1])$ for which $f = 1$ on K, and $f = 0$ outside a compact subset of O.*

Proof. Since K is compact, by Proposition 7.2.4 there is a precompact open V such that $K \subseteq V \subseteq \overline{V} \subseteq O$. Now $\overline{V}$ is compact and hence normal. Since K and $\overline{V} - V$ are disjoint sets closed in $\overline{V}$, by Urysohn's Lemma 7.1.10 there is a $f \in C(\overline{V},[0,1])$ such that $f|_K \equiv 1$ and $f|_{\overline{V}-V} \equiv 0$. We extend f on X by setting $f|_{X-\overline{V}} \equiv 0$. Let E be a closed subset of $[0,1]$. If $0 \notin E$, then $f^{-1}(E) = (f|_{\overline{V}})^{-1}(E)$ is closed by continuity on $\overline{V}$. If $0 \in E$, then $f^{-1}(E) = (f|_{\overline{V}})^{-1}(E) \cup (X - \overline{V}) = (f|_{\overline{V}})^{-1}(E) \cup (X - V)$, the last equality holds because $(f|_{\overline{V}})^{-1}(E) \supseteq \partial V$, hence $f^{-1}(E)$ is also closed, continuity follows. □

The author in [**?**] introduced a useful notation for functions constructed in LCH version of Urysohn's lemma. The notation
$$K \prec f$$

means K is a compact subset of X and $f \in C_c(X,[0,1])$, $f|_K \equiv 1$. The notation

$$f \prec O$$

means that O is open, $f \in C_c(X,[0,1])$ and $\operatorname{spt} f \subseteq O$.

Adopting the above notation, Lemma 7.2.5 can be rewritten as:

Lemma 7.2.6 (Urysohn, LCH Version). *Let X be an LCH space. For every compact subset K of X and every open $O \supseteq K$, there is f, $K \prec f \prec O$.*

Theorem 7.2.7 (Tietze Extension, LCH Version). *Let X be an LCH space and K compact subset of X. If $f \in C(K)$, then there is $F \in C_c(X)$ such that $F|_K = f$.*

As in the situation in normal space, by applying LCH version of Urysohn's lemma we can derive LCH version of Tietze extension theorem. This time will be easier due to the earlier work.

Proof. Since X contains K, K has a precompact neighborhood V. Since $\overline{V}$ is compact, $\overline{V}$ has a precompact neighborhood U:

$$K \subseteq V \subseteq \overline{V} \subseteq U \subseteq \overline{U}.$$

Since $\overline{U}$ is a normal space and K is closed in $\overline{U}$, by Theorem 7.1.17 there is $F \in C(\overline{U})$ such that $F|_K = f$. By Lemma 7.1.10 there is $\varphi \in C_c(X,[0,1])$ such that $K \prec \varphi \prec V$.

Now we show that $G := F\varphi$ is desired extension. Firstly, $G|_K = F|_K \varphi|_K = f$, and secondly, both $G|_{X-\overline{V}} = 0$ and $G|_U$ are continuous with $(X - \overline{V}) \cup U = X$, hence G is continuous on X. □

Given a topological space X and a set $E \subseteq X$. A **partition of unity** on E is a collection

$$\{h_\alpha \in C(X,[0,1])\}_{\alpha \in A}$$

of functions satisfying:

(i) Each $x \in X$ has a neighborhood such that h_α's are zero except finitely many of them.

(ii) $\sum_{\alpha \in A} h_\alpha = 1$ on E.

Moreover, a partition of unity $\{h_\alpha\}$ is subordinate to an open cover $\mathcal{U}$ of E if for each α, there is $U \in \mathcal{U}$ such that $\operatorname{spt} h_\alpha \subseteq U$.

Proposition 7.2.8. *Let X be an LCH space, K a compact subset of X and $\{U_j\}_{j=1}^n$ an open covering of K. There is a partition of unity on K subordinate to $\{U_j\}_{j=1}^n$ consisting of compactly supported functions.*

Proof. Let $x \in K$, then there is an i and a precompact neighborhood O_x of x such that $\overline{O_x}$ is contained in U_i. Since K is compact, there are $x_1, x_2, \dots, x_n \in K$ such that $K \subseteq \bigcup_{i=1}^n O_{x_i}$. For $j = 1,2,\dots,n$, define $F_j = \bigcup_{\overline{O_{x_i}} \subseteq U_j} \overline{O_{x_i}}$. F_j is compact (probably empty) and $F_j \subseteq U_j$, hence there is $g_j \in C_c(X,[0,1])$ such that $F_j \prec g_j \prec U_j$.

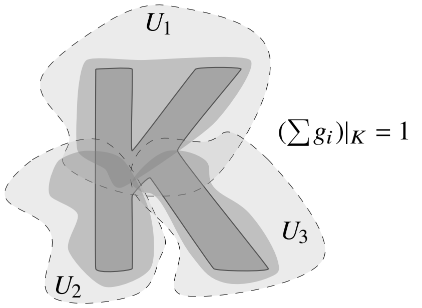


Figure 7.2: Partition of unity on K.

Since the open set $\{\sum_{i=1}^n g_i > 0\} \supseteq K$, there is $f \in C_c(X,[0,1])$ such that $K \prec f \prec \{\sum_{i=1}^n g_i > 0\}$. Hence if $\sum_{i=1}^n g_i(x) = 0$, $f(x) = 0$ and if $x \in K$, $f(x) = 1$, so the functions

$$h_i := \frac{g_i}{g_1 + \cdots + g_n + (1 - f)},$$

$i = 1, 2, \ldots, n$, forms a partition of unity on K subordinate to $\{U_i\}_{i=1}^n$. $\square$

Remark. A useful form of Proposition 7.2.8 is the following: Let K be compact and $\{U_i\}_{i=1}^n$ an open cover of K in a LCH space X, then there are $g_1, \ldots, g_n \in C_c(X,[0,1])$ such that $g_i \prec U_i$ and $\sum g_i = 1$ on K.

7.3 Riesz Representation Theorem and Regularity of Measures

In this section let's denote X a LCH space.

Definition 7.3.1. Let μ be a Borel measure on X. μ is said to be **outer regular** if for every Borel set E,

$$\mu(E) = \inf\{\mu(U) : U \supseteq E, U \text{ open}\}$$

and **inner regular** if for every Borel set E,

$$\mu(E) = \sup\{\mu(K) : K \subseteq E, K \text{ compact}\}.$$

μ is said to be **regular** if μ is both outer and inner regular. Finally, a Borel measure μ is **Radon** if μ is outer regular for Borel sets, inner regular for open sets and finite for compact sets.

Remark. It can be shown that if μ is a Radon measure on X and X is σ-finite, then μ is regular. See Proposition 7.3.7 for detail.

Definition 7.3.2. A linear functional Λ on $C_c(X)$ is said to be **positive** if $\Lambda(f) \geq 0$ whenever $f \geq 0$.

Theorem 7.3.3 (Riesz Representation Theorem). *If Λ is a positive linear functional on $C_c(X)$, then there is a unique Radon measure μ on X such that $\Lambda(f) = \int_X f\, d\mu$ for all $f \in C_c(X)$.*

For open set U in X we can assign it a nonnegative value

$$\mu_0(U) = \sup\{\Lambda(f) : f \prec U\} \tag{7.3.4}$$

and for every set $E \subset X$ we define

$$\mu(E) = \inf\{\mu_0(U) : U \supseteq E, U \text{ open}\} \tag{7.3.5}$$

and define $\mu(\emptyset) = 0$. Let U, V be open, $U \subseteq V$, then $\mu_0(U) \leq \mu_0(V)$, and hence we can prove that $\mu_0(U) = \mu(U)$. In the following proof we define μ to be the one in (7.3.5) and we don't distinguish μ and μ_0 for open sets.

Proof. Our aim is to prove μ defined in (7.3.5) is the desired Radon measure, the uniqueness part will be a simple application of Urysohn's lemma and proved in the last step of the proof.

Step 1 (μ is an outer measure). Let $E \subseteq \bigcup_{i=1}^\infty A_i$, we need to show $\mu(E) \leq \sum_{i=1}^\infty \mu(A_i)$. We may assume $\mu(A_i) < \infty$. Let $\epsilon > 0$ be given, then by definition in (7.3.5) there is an open set $U_i \supseteq A_i$ such that

$$\mu(U_i) < \mu(A_i) + \frac{\epsilon}{2^i}.$$

Now $\bigcup_{i=1}^\infty U_i \supseteq E$ is open, we have $\mu(E) \leq \mu(\bigcup_{i=1}^\infty U_i)$ by definition (7.3.5). Next to further estimate $\mu(\bigcup_{i=1}^\infty U_i)$ we need to use (7.3.4). Let $f \prec \bigcup_{i=1}^\infty U_i$, then by compactness there is an n such that $\operatorname{spt} f \subseteq \bigcup_{i=1}^n U_i$. By Proposition 7.2.8 there is $\varphi_1, \varphi_2, \ldots, \varphi_n$ such that $\varphi_i \prec U_i$ and $(\sum_{i=1}^n \varphi_i)|_{\operatorname{spt} f} = 1$. Now

$$f = \sum_{i=1}^n f\varphi_i \implies \Lambda f = \sum_{i=1}^n \Lambda(f\varphi_i) \leq \sum_{i=1}^n \mu(U_i) \leq \sum_{i=1}^\infty \mu(U_i) \leq \sum_{i=1}^\infty \mu(A_i) + \epsilon.$$

Since this is true for each $f \prec \bigcup_{i=1}^\infty U_i$, by taking supremum we have

$$\mu\left(\bigcup_{i=1}^\infty U_i\right) \leq \sum_{i=1}^\infty \mu(A_i) + \epsilon,$$

and this is true for each fixed $\epsilon > 0$, hence step 1 is completed.

Step 2 (μ is a Borel measure). It is enough to show every open set G satisfies

$$\mu(E) \geq \mu(E \cap G) + \mu(E - G). \tag{7.3.6}$$

for every set $E \subseteq X$. To prove this, it is enough to prove (7.3.6) holds when E is open[(3)]. Let's assume E is open. In view of definition in (7.3.4) let's fix an $f \prec E \cap G$. Then fix $g \prec E - \operatorname{spt} f$, we have $f + g \prec E$, and hence

$$\mu(E) \geq \Lambda(f + g) = \Lambda f + \Lambda g,$$

this is true for every $g \prec E - \operatorname{spt} f$, hence

$$\mu(E) \geq \Lambda f + \mu(E - \operatorname{spt} f) \geq \Lambda f + \mu(E - G).$$

[(3)]It is because after that for any set $F \subseteq X$, we can let $O \supseteq F$ be open and

$$\mu(O) \geq \mu(O \cap G) + \mu(O - G) \geq \mu(F \cap G) + \mu(F - G).$$

By taking infimum, (7.3.5) tells us $\mu(F) \geq \mu(F \cap G) + \mu(F - G)$, so that G is measurable.

Now this is true for every $f \prec E \cap G$, hence

$$\mu(E) \geq \mu(E \cap G) + \mu(E - G).$$

Step 3 ($\boldsymbol{\mu}$ is finite on compact sets). This is a consequence of the following formula for compact set K:

$$\mu(K) = \inf\{\Lambda f : f \succ K\}.$$

Let K be compact and $U \supseteq K$ open, then by Urysohn's lemma we can find an f such that $U \succ f \succ K$, so

$$\mu(U) \geq \Lambda f \geq \inf\{\Lambda f : f \succ K\}.$$

This is true for each open $U \supseteq K$, so by (7.3.5) $\mu(K) \geq \inf\{\Lambda f : f \succ K\}$.

Conversely, fix an $f \succ K$, then we fix an $\alpha \in (0,1)$, then the set $G_\alpha = \{f > \alpha\}$ is open and $G_\alpha \supseteq K$. For every $g \prec G_\alpha$, we have $g \leq f/\alpha$, hence

$$\mu(K) \leq \mu(G_\alpha) = \sup\{\Lambda g : g \prec G_\alpha\} \leq \frac{\Lambda f}{\alpha}.$$

As this is true for every $\alpha \in (0,1)$, we have $\mu(K) \leq \Lambda f$. This is true for every $f \succ K$, $\mu(K) \leq \inf\{\Lambda f : f \succ K\}$, completing step 3.

Step 4 ($\boldsymbol{\Lambda f = \int_X f\, d\mu}$). To do this, it is enough to show $\Lambda f \leq \int_X f\, d\mu$ for every $f \in C_c(X)$. Let $f \in C_c(X)$ be given, then let $f(X) = [a,b]$. Let's fix an $\epsilon > 0$ and let $y_0, y_1, y_2, \ldots, y_n$ be such that

$$y_0 < a < y_1 < y_2 < \cdots < y_n = b$$

with $|y_i - y_{i-1}| < \epsilon$ for each $i = 1,2,\ldots,n$. We let $G_i = f^{-1}(y_{i-1}, y_i] \cap \operatorname{spt} f$, then $\operatorname{spt} f = \bigsqcup_{i=1}^n G_i$. Each G_i is Borel measurable with finite measure, by (7.3.5) for each i we can find an open set $U_i \supseteq G_i$ such that $\mu(U_i) < \mu(G_i) + \epsilon/n$. We can further assume $f(U_i) \subseteq (y_{i-1}, y_i + \epsilon)$. Now

$$\operatorname{spt} f \subseteq \bigcup_{i=1}^n U_i,$$

we can find a partition of unity $\{\varphi_1, \varphi_2, \ldots, \varphi_n\}$ of $\operatorname{spt} f$ subordinate to $\{U_i\}$ with $\varphi_i \prec U_i$. Now $f = \sum_{i=1}^n f\varphi_i$ and hence

$$\begin{aligned}
\Lambda f &= \sum_{i=1}^n \Lambda(f\varphi_i) \leq \sum_{i=1}^n \Lambda((y_i + \epsilon)\varphi_i) \\
&= \sum_{i=1}^n (y_i + \epsilon)\Lambda\varphi_i \\
&= \sum_{i=1}^n \underbrace{(|a| + y_i + \epsilon)}_{>0}\Lambda\varphi_i - |a| \sum_{i=1}^n \Lambda\varphi_i \\
&\leq \sum_{i=1}^n (|a| + y_i + \epsilon)\mu(U_i) - |a|\Lambda\left(\sum_{i=1}^n \varphi_i\right) \\
&\leq \sum_{i=1}^n (|a| + y_{i-1} + 2\epsilon)\left(\mu(G_i) + \frac{\epsilon}{n}\right) - |a|\Lambda\left(\sum_{i=1}^n \varphi_i\right)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n}(|a|+y_{i-1}+2\epsilon)\frac{\epsilon}{n}+\sum_{i=1}^{n}(|a|+y_{i-1}+2\epsilon)\mu(G_i)-|a|\Lambda\left(\sum_{i=1}^{n}\varphi_i\right)\\
&\le (|a|+b+2\epsilon)\epsilon+|a|\mu(\operatorname{spt} f)+\sum_{i=1}^{n}y_{i-1}\mu(G_i)+2\epsilon\mu(\operatorname{spt} f)-|a|\Lambda\left(\sum_{i=1}^{n}\varphi_i\right)\\
&\le (|a|+b+2\epsilon)\epsilon+\int_X f\,d\mu+2\epsilon\mu(\operatorname{spt} f)+|a|\underbrace{\left(\mu(\operatorname{spt} f)-\Lambda\left(\sum_{i=1}^{n}\varphi_i\right)\right)}_{\le 0 \text{ by step 3}}\\
&\le (|a|+b+2\epsilon)\epsilon+2\epsilon\mu(\operatorname{spt} f)+\int_X f\,d\mu.
\end{aligned}
$$

We let $\epsilon \to 0$ to conclude $\Lambda f \le \int_X f\,d\mu$.

Step 5 (μ is Radon). μ is outer regular for Borel sets by definition in (7.3.5). In step 3 we have shown that μ is finite for compact sets. It remains to show μ is inner regular for open sets. Recall by definition in (7.3.4) we have for open set U,

$$\mu(U)=\sup\{\Lambda f : f \prec U\}=\sup\left\{\int_{\operatorname{spt} f} f\,d\mu : f \prec U\right\}\le \sup\{\mu(\operatorname{spt} f) : f \prec U\},$$

hence $\mu(U)=\sup\{\mu(\operatorname{spt} f) : f \prec U\}$, so μ is inner regular.

Step 6 (Such μ is unique). Suppose there is a Radon measure ν such that $\int_X f\,d\mu=\int_X f\,d\nu$ for every $f\in C_c(X)$, we try to show $\mu=\nu$. In fact, let $\epsilon>0$ be given and let K be compact, then there is an open $U\supseteq K$ such that $\mu(U)-\mu(K), \nu(U)-\nu(K)<\epsilon$. By Urysohn's lemma there is an f such that $K\prec f\prec U$, now $\int_X f\,d\mu=\int_X f\,d\nu \implies \int_U f\,d\mu=\int_U f\,d\nu$, and hence

$$|\mu(K)-\nu(K)|\le\left|\int_{U-K} f\,d\mu\right|+\left|\int_{U-K} f\,d\nu\right|<2\epsilon.$$

Since $\epsilon>0$ can be fixed arbitrarily, we have $\mu(K)=\nu(K)$. Since μ and ν are inner regular for open sets, we have $\mu(U)=\nu(U)$ for every open set U. Finally since μ and ν are outer regular, we have $\mu(E)=\nu(E)$ for every Borel set E. □

Proposition 7.3.7. *Let μ be a Radon measure on X.*

(i) If $\mu(E)<\infty$, then E is inner regular.

(ii) If X is σ-finite, then μ is a regular measure on X. Moreover, for every Borel set E and fixed $\epsilon>0$, there is an open U and a closed L with $L\subseteq E\subseteq U$ and $\mu(U-L)<\epsilon$.

Proof. (i) Let $\epsilon>0$ be given. By outer regularity we can find an open $U\supseteq E$ such that $\mu(U)-\mu(E)<1$[(4)]. Let $V\supseteq U-E$ be such that $\mu(V-(U-E))<\epsilon$. Now we expect $U-V\subseteq E$ is a nice approximation. Since $U-V$ is not compact, we further choose a compact $K\subseteq U$ such that $\mu(U)-\mu(K)<\epsilon$. Now the compact set $K-V\subseteq E$ should be good enough. Since

$$E-(K-V)=(E-K)\cup(E\cap V)$$

[(4)]We are trying to approximate $U-E$ from outside to get inner approximation of E, so the approximation U of E needs not be tight.

$$= (E-K)\cup(V-(V-E))$$
$$\subseteq (U-K)\cup(V-(U-E)),$$

so we have

$$\mu(E-(K-V)) \le \mu(U-K)+\mu(V-(U-E)) < \epsilon+\epsilon = 2\epsilon.$$

(ii) In view of (i) to show μ is inner regular, it is enough to show when $\mu(E)=\infty$, we have $\sup\{\mu(K) : K \subseteq E, K \text{ compact}\} = \infty$. Suppose $\mu(E)=\infty$. Let $X = \bigsqcup_{i=1}^{\infty} X_i$ with $\mu(X_i)<\infty$. Then $\mu(E)=\sum_{i=1}^{\infty}\mu(E\cap X_i)$. So for every $b>1$, there is an n such that

$$\sum_{i=1}^{n}\mu(E\cap X_i) > b.$$

By part (i) we can find a compact $K_i \subseteq E\cap X_i$ such that $\mu(E\cap X_i)-\mu(K_i) < \frac{1}{2^i}$, now $K := \bigcup_{i=1}^{n} K_i$ is compact and $\mu(K) > b-1$, as desired.

Let $\epsilon>0$ be given, we try to prove the second statement in part (ii). Since E is σ-finite, $E=\bigcup_{i=1}^{\infty}E_i$ with $\mu(E_i)<\infty$. As μ is outer regular, we can find an open $U_i \supseteq E_i$ such that $\mu(U_i)-\mu(E_i) < \epsilon/2^i$. Let $U=\bigcup_{i=1}^{\infty}U_i$ we have

$$\mu(U-E) \le \sum_{i=1}^{\infty}\mu(U_i-E) \le \sum_{i=1}^{\infty}\mu(U_i-E_i) < \epsilon.$$

Similarly we can find an open set $V\supseteq E^c$ such that $\mu(V-E^c)<\epsilon$. Hence

$$\mu(U-V^c) \le \mu(U-E)+\mu(\underbrace{E-V^c}_{=V-E^c}) < \epsilon+\epsilon = 2\epsilon.$$

Note that $L = V^c \subseteq E$ is closed. □

Definition 7.3.8. A set A in a topological space Y is σ-compact if A is a countable union of compact sets in Y.

A useful case is that every open set in a second countable LCH space is σ-compact.

Theorem 7.3.9. *If every open set in X is σ-compact, then every Borel measure on X that is finite on compact sets is regular.*

Proof. Let μ be a Borel measure on X that is finite on compact sets, then $\Lambda(f) := \int_X f\,d\mu$ defines a positive functional on $C_c(X)$, hence there is a Radon measure ν on X such that $\int_X f\,d\mu = \int_X f\,d\nu$ for every $f\in C_c(X)$. Let U be open, then $U=\bigcup_{i=1}^{\infty}K_i$ for some compact sets K_i. Let $K_1 \prec f_1 \prec U$, and for each $n\ge 2$ we construct

$$\bigcup_{i=1}^{n}K_i \cup \left(\bigcup_{i=1}^{n-1}\operatorname{spt} f_i\right) \prec f_n \prec U,$$

then $\{f_n\}$ is pointwise increasing and $\lim_{n\to\infty} f_n(x) = \chi_U(x)$. Now

$$\mu(U) = \lim_{n\to\infty}\int_X f_n\,d\mu = \lim_{n\to\infty}\int_X f_n\,d\nu = \nu(U). \tag{7.3.10}$$

So μ and ν agree on open sets[5].

Since ν is Radon and X is σ-compact, X is σ-finite w.r.t. ν. Let E be Borel and $\epsilon > 0$ be given, by Proposition 7.3.7 we can find open V and closed L such that $L \subseteq E \subseteq V$ and $\nu(V - L) < \epsilon$. Since $V - L$ is open, by (7.3.10) we have $\mu(V - L) < \epsilon$ and therefore

$$\mu(V) \leq \mu(E) + \epsilon, \quad \mu(E) \leq \mu(L) + \epsilon.$$

The first inequality shows that E is outer regular. Since there are compact sets L_i such that $L = \bigcup_{i=1}^{\infty} L_i$, we have $\mu(L) = \lim_{n\to\infty} \mu(\bigcup_{i=1}^{n} L_i)$, so μ is also inner regular. □

[5] We are not yet done because μ is not necessarily outer regular.