# Real Analysis, Probability, Random Processes

with Measure Theory

Written by Jaeyoung Lee

Updated on Feb., 9th, 2022.

# Contents

# Chapter 1

# $\sigma$-Algebra and Measurable Functions

In this part, we summarize some of the basis theory on $\sigma$-algebra and measurable functions, which have served as the fundamentals to measure theory, Lebesque integration, probability, etc. The materials either directly come from or is strongly related to those in Chapters 1.2, 2.1, and 2.4 in Folland (1999)'s real analysis book. Related to them, this note also introduces some basic theory of topology and set theory that also make the statements about $\sigma$-algebra and measurable functions clear and rigorous.

## 1.1 $\sigma$-algebra

$$\text{In this note, } \begin{cases} X \text{ and } Y \text{ denote any non-empty sets;} \\ f: X \to Y \text{ is any mapping from } X \text{ to } Y; \\ \mathcal{M} \subseteq \mathcal{P}(X) \text{ and } \mathcal{N} \subseteq \mathcal{P}(Y) \text{ denote any (or some given) } \sigma\text{-algebras on } X \text{ and } Y\text{, respectively.} \end{cases}$$

$(X, \mathcal{M})$ (and $(Y, \mathcal{N})$) is called a measurable space. Here, a $\sigma$-algebra $\mathcal{M}$ on $X$ (or $\mathcal{N}$ on $Y$) is precisely defined as follows.

**Definition 1.1.** *A non-empty subset $\mathcal{M} \subseteq \mathcal{P}(X)$ is said to be a $\sigma$-algebra on $X$ iff:*

$$\text{1) } \{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M} \implies \bigcup_{i=1}^{\infty} E_i \in \mathcal{M} \qquad \textit{(closed under countable unions)},$$

$$\text{2) } E \in \mathcal{M} \implies E^c \in \mathcal{M} \qquad \textit{(closed under complements)}.$$

*If the first property is replaced by*

$$\text{1}'\text{) } \{E_i\}_{i=1}^{N} \subseteq \mathcal{M} \implies \bigcup_{i=1}^{N} E_i \in \mathcal{M} \qquad \textit{(closed under finite unions)},$$

*then, $\mathcal{M}$ is said to be an algebra on $X$.*

Note that any $\sigma$-algebra on $X$ is an algebra on $X$, but not vice versa. As discussed later, a $\sigma$-algebra and an algebra are the domains of a measure and a premeasure, respectively. The basic properties of a $\sigma$-algebra (and an algebra) are as follows.

**Proposition 1.1.** *The followings hold for any $\sigma$-algebra $\mathcal{M}$ or any algebra $\mathcal{M}$ on $X$:*

1. $\varnothing$, $X \in \mathcal{M}$,

$$\text{2. } \begin{cases} \text{for a } \sigma\text{-algebra } \mathcal{M}\text{: } \{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M} \implies \bigcap_{i=1}^{\infty} E_i \in \mathcal{M} & \textit{(closed under countable intersections)}, \\ \text{for an algebra } \mathcal{M}\text{: } \{E_i\}_{i=1}^{N} \subseteq \mathcal{M} \implies \bigcap_{i=1}^{N} E_i \in \mathcal{M} & \textit{(closed under finite intersections)}, \end{cases}$$

*3.* $E,\ F \in \mathcal{M} \Longrightarrow E \setminus F \in \mathcal{M}$ *and* $E \triangle F \in \mathcal{M}$.

*Proof.* The second part can be directly proven as follows:

$$\text{for a } \sigma\text{-algebra } \mathcal{M}\text{: } \bigcap_{i=1}^{\infty} E_i = \left( \bigcup_{i=1}^{\infty} E_i^c \right)^c \in \mathcal{M}, \quad \left( \text{for an algebra } \mathcal{M}\text{: } \bigcap_{i=1}^{N} E_i = \left( \bigcup_{i=1}^{N} E_i^c \right)^c \in \mathcal{M} \right).$$

Then, the first part is obvious since for $E \in \mathcal{M}$, we have $X = E \cup E^c \in \mathcal{M}$ and $\varnothing = E \cap E^c \in \mathcal{M}$. Likewise, since $\mathcal{M}$ is closed under countable unions, countable intersections, and complements, we finally have $E \setminus F = E \cap F^c \in \mathcal{M}$ and $E \triangle F \doteq (E \setminus F) \cup (F \setminus E) \in \mathcal{M}$. □

**Example 1.1.** $\mathcal{M} = \mathcal{P}(X)$ *and* $\mathcal{M} = \{X, \varnothing\}$ *are the largest and smallest $\sigma$-algebra on $X$, respectively.*

The following shows that the (uncountable) intersection of $\sigma$-algebras is also a $\sigma$-algebra.

**Proposition 1.2.** *Let $\{\mathcal{M}_\alpha\}_{\alpha \in A}$ be a family of $\sigma$-algebras on $X$. Then, $\mathcal{M} = \bigcap_{\alpha \in A} \mathcal{M}_\alpha$ is a $\sigma$-algebra.*

*Proof.* **(closed under complements)** Assume $E \in \mathcal{M}$. Then, by definition, $E \in \mathcal{M}_\alpha$ for all $\alpha \in A$, which again implies $E^c \in \mathcal{M}_\alpha$ for all $\alpha \in A$. Hence, we obtain $E^c \in \bigcap_{\alpha \in A} \mathcal{M}_\alpha = \mathcal{M}$.

**(closed under countable unions)** Assume $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$. Then, $E_i \in \mathcal{M}$ implies that $E_i \in \mathcal{M}_\alpha$ for all $\alpha \in A$. Since $\mathcal{M}_\alpha$ is a $\sigma$-algebra, we obtain $\bigcup_{i=1}^{\infty} E_i \in \mathcal{M}_\alpha$ for all $\alpha \in A$, implying that $\bigcup_{i=1}^{\infty} E_i \in \mathcal{M}$. □

**Definition 1.2.** *For any family $\mathcal{E} \subseteq \mathcal{P}(X)$, $\sigma(\mathcal{E})$ denotes the smallest $\sigma$-algebra on $X$ that contains $\mathcal{E}$; we call $\sigma(\mathcal{E})$ the $\sigma$-algebra generated by $\mathcal{E}$.*

**Remark 1.1.** *Note that for any family $\mathcal{E} \subseteq \mathcal{P}(X)$, there is at least one $\sigma$-algebra, namely, the power set $\mathcal{P}(X)$ itself, the largest $\sigma$-algebra, that contains $\mathcal{E}$. Moreover, since any (uncountable) intersection of $\sigma$-algebras is also a $\sigma$-algebra as shown in Proposition 1.2, the smallest $\sigma$-algebra $\sigma(\mathcal{E})$ in Definition B.1 can be constructed and recognized as the intersection of all $\sigma$-algebras containing $\mathcal{E}$.*

By Proposition 1.2, the following lemma is obvious.

**Lemma 1.1.** *If $\mathcal{E} \subseteq \mathcal{M}$, then $\sigma(\mathcal{E}) \subseteq \mathcal{M}$.*

## 1.2 Abstract Measurable Functions

**Lemma 1.2.** *The inverse mapping $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ defined by $f^{-1}(E) \doteq \{x \in X : f(x) \in E\}$ preserves the unions, intersections, and complements. That is,*

$$\text{1) } f^{-1}\left( \bigcup_{\alpha \in A} E_\alpha \right) = \bigcup_{\alpha \in A} f^{-1}(E_\alpha),$$
$$\text{2) } f^{-1}\left( \bigcap_{\alpha \in A} E_\alpha \right) = \bigcap_{\alpha \in A} f^{-1}(E_\alpha),$$
$$\text{3) } f^{-1}(E^c) = (f^{-1}(E))^c.$$

*for any indexed family $\{E_\alpha\}_{\alpha \in A} \subseteq \mathcal{P}(Y)$ and any $E \in \mathcal{P}(Y)$.*

**Proposition 1.3.** *Any $\mathcal{M} \subseteq \mathcal{P}(X)$ given by $\mathcal{M} = \{f^{-1}(E) : E \in \mathcal{N}\}$ is a $\sigma$-algebra on $X$.*

*Proof.* Let $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{N}$ and $E \in \mathcal{N}$. Then, $\bigcup_{i=1}^{\infty} E_i \in \mathcal{N}$ and $E^c \in \mathcal{N}$ since $\mathcal{N}$ is a $\sigma$-algebra. Hence,

$$\bigcup_{i=1}^{\infty} f^{-1}(E_i) = f^{-1}\left( \bigcup_{i=1}^{\infty} E_i \right) \in \mathcal{M}$$
$$(f^{-1}(E))^c = f^{-1}(E^c) \in \mathcal{M}$$

by Lemma 1.2. □

**Definition 1.3.** *A function $f : X \to Y$ is said to be $(\mathcal{M}, \mathcal{N})$-measurable iff:* $E \in \mathcal{N} \implies f^{-1}(E) \in \mathcal{M}$.

**Proposition 1.4.** *A function $f : X \to Y$ is always $(\mathcal{M}, \mathcal{N})$-measurable for $\mathcal{M}$ given in Proposition 1.3,*

*Proof.* Trivial by Definition 1.3. □

The $(\mathcal{M}, \mathcal{N})$-measurability can be understood in a sense that the restricted function $f^{-1}|_{\mathcal{N}} : \mathcal{N} \to \mathcal{M}$ on the $\sigma$-algebra domain $\mathcal{N} \subseteq \mathcal{P}(Y)$ is well-defined in such a way that its image is contained by its codomain $\mathcal{M}$ as in the usual definition of a function. This observation can be summarized as follows.

**Proposition 1.5.** *$f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable iff* $\operatorname{Im}(f^{-1}|_{\mathcal{N}}) \subseteq \mathcal{M}$.

**Corollary 1.1.** *If $\overline{\mathcal{M}}$ and $\underline{\mathcal{N}}$ are respective $\sigma$-algebras on $X$ and $Y$ such that $\mathcal{M} \subseteq \overline{\mathcal{M}}$ and $\underline{\mathcal{N}} \subseteq \mathcal{N}$, then*

$$f \text{ is } (\mathcal{M}, \mathcal{N})\text{-measurable} \implies f \text{ is } (\overline{\mathcal{M}}, \underline{\mathcal{N}})\text{-measurable.}$$

*Proof.* Trivial by Proposition 1.5 and $\operatorname{Im}(f^{-1}|_{\underline{\mathcal{N}}}) \subseteq \operatorname{Im}(f^{-1}|_{\mathcal{N}}) \subseteq \mathcal{M} \subseteq \overline{\mathcal{M}}$. □

**Corollary 1.2.** *For any $\sigma$-algebra $\mathcal{N}$ on $Y$, any function $f : X \to Y$ is $(\mathcal{P}(X), \mathcal{N})$-measurable.*

*Proof.* Trivial by $\operatorname{Im}(f^{-1}|_{\mathcal{N}}) \subseteq \mathcal{P}(X)$, where $\mathcal{P}(X)$ is the largest $\sigma$-algebra on $X$ (see Example 1.1). □

The following is the dual to Propositions 1.3 and 1.4.

**Lemma 1.3.** *Any $\mathcal{N} \subseteq \mathcal{P}(Y)$ given by $\mathcal{N} = \{E : f^{-1}(E) \in \mathcal{M}\}$ is a $\sigma$-algebra on $Y$; for such $\mathcal{N}$, $f$ is always $(\mathcal{M}, \mathcal{N})$-measurable.*

*Proof.* Let $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{N}$ and $E \in \mathcal{N}$. Then, we have $f^{-1}(E_i) \subseteq \mathcal{M}$ for each $i \in \mathbb{N}$ and $f^{-1}(E) \in \mathcal{M}$. Since $\mathcal{M}$ is a $\sigma$-algebra, we therefore obtain by Lemma 1.2

$$f^{-1}\left(\bigcup_{i=1}^{\infty} E_i\right) = \bigcup_{i=1}^{\infty} f^{-1}(E_i) \in \mathcal{M},$$
$$f^{-1}(E^c) = (f^{-1}(E))^c \in \mathcal{M},$$

meaning that $\bigcup_{i=1}^{\infty} E_i \in \mathcal{N}$ and $E^c \in \mathcal{N}$. Therefore, $\mathcal{N}$ is a $\sigma$-algebra on $Y$. Moreover, the definition of $\mathcal{N}$ gives $\operatorname{Im}(f^{-1}|_{\mathcal{N}}) \subseteq \mathcal{M}$, and hence $f$ is $(\mathcal{M}, \mathcal{N})$-measurable by Proposition 1.5. □

**Proposition 1.6.** *For any subset $\mathcal{E} \subseteq Y$, $f$ is $(\mathcal{M}, \sigma(\mathcal{E}))$-measurable iff* $\operatorname{Im}(f^{-1}|_{\mathcal{E}}) \subseteq \mathcal{M}$, *i.e., iff*

$$f^{-1}(E) \in \mathcal{M} \text{ for all } E \in \mathcal{E}.$$

*Proof.* ($\Longrightarrow$) If $f$ is $(\mathcal{M}, \sigma(\mathcal{E}))$-measurable, then $\mathcal{E} \subseteq \sigma(\mathcal{E})$ and Proposition 1.5 imply that

$$\operatorname{Im}(f^{-1}|_{\mathcal{E}}) \subseteq \operatorname{Im}(f^{-1}|_{\sigma(\mathcal{E})}) \subseteq \mathcal{M}.$$

($\Longleftarrow$) The assumption assures the expression $\mathcal{E} = \{E \in \mathcal{E} : f^{-1}(E) \in \mathcal{M}\}$. Moreover, the $\sigma$-algebra $\mathcal{N}$ given in Lemma 1.3 obviously satisfies $\mathcal{E} = \{E \in \mathcal{E} : f^{-1}(E) \in \mathcal{M}\} \subseteq \{E \in \mathcal{P}(Y) : f^{-1}(E) \in \mathcal{M}\} = \mathcal{N}$ and hence, $\sigma(\mathcal{E}) \subseteq \mathcal{N}$ by Lemma 1.1. This obviously results in $\operatorname{Im}(f^{-1}|_{\sigma(\mathcal{E})}) \subseteq \operatorname{Im}(f^{-1}|_{\mathcal{N}}) \subseteq \mathcal{M}$, and hence $f$ is $(\mathcal{M}, \sigma(\mathcal{E}))$-measurable by Proposition 1.5. □

**Proposition 1.7.** *Suppose $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable and $g : Y \to Z$ is $(\mathcal{N}, \mathcal{O})$-measurable. Then, the composition $g \circ f : X \to Z$ is $(\mathcal{M}, \mathcal{O})$-measurable.*

*Proof.* By measurability, $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{N}$ and $g^{-1}(F) \in \mathcal{N}$ for all $F \in \mathcal{O}$. Hence, it is obvious that $f^{-1}(g^{-1}(F)) \in \mathcal{M}$ for all $F \in \mathcal{O}$, and the proof is completed by $(g \circ f)^{-1} = f^{-1} \circ g^{-1} = f^{-1}(g^{-1}(\cdot))$. □

## 1.3 Product $\sigma$-Algebra and Measurable Functions on a Cartesian Product

Let $\{Y_\alpha\}_{\alpha\in A}$ be a family of non-empty sets $Y_\alpha$ indexed by $\alpha \in A$.

**Definition 1.4.** *The Cartesian product $\prod_{\alpha\in A} Y_\alpha$ is the set of all functions $y : A \to \bigcup_{\alpha\in A} Y_\alpha$ such that $y(\alpha) \in Y_\alpha$ for every $\alpha \in A$. We denote $y_\alpha \doteq y(\alpha)$ and call it the $\alpha$-th element of $y$.*

If the sets $Y_\alpha$ are all equal to some fixed set $Z$, then we denote $\prod_{\alpha\in A} Y_\alpha$ by $Z^A$. Moreover, if

$$A \text{ is finite and given by } A = \{1, 2, 3, \cdots, n\} \text{ for some } n \in \mathbb{N}, \tag{1.1}$$

then $\prod_{\alpha\in A} Y_\alpha$ and $Z^A$ will be denoted by $Y_1 \times Y_2 \times \cdots \times Y_n$ and $Z^n$, respectively. For notational convenience, we also denote $Y \doteq \prod_{\alpha\in A} Y_\alpha$ throughout this section.

**Definition 1.5.** *The $\alpha$-th projection map $p_\alpha : Y \to Y_\alpha$ is such that $p_\alpha(y) = y_\alpha$ for all $y \in Y$.*

**Remark 1.2.** *When (1.1) is true for the indexed set $A$, the Cartesian product $Y = Y_1 \times Y_2 \times \cdots \times Y_n$ is usually defined as the set of all tuples $(y_1, y_2, \cdots, y_n)$ such that $y_\alpha \in Y_\alpha$ for $\alpha = 1, 2, \cdots, n$. This definition is set-theoretically different from the Cartesian product defined above in Definition 1.4. However, both definitions with the notation $y_\alpha \doteq y(\alpha)$ above agree that*

$$y_\alpha \in Y_\alpha \text{ for any } \alpha \in A \text{ and thus } p_\alpha(y) = y_\alpha \text{ for any } \alpha \in A \text{ and all } y \in Y. \tag{1.2}$$

*Likewise, all of the theorems and proofs will be also true for the finite case (1.1).*

Now, let for each $\alpha \in A$: $\begin{cases} \mathcal{N}_\alpha \subseteq \mathcal{P}(Y_\alpha) \text{ be a } \sigma\text{-algebra on } Y_\alpha; \\ f_\alpha : X \to Y_\alpha \text{ be a function from } X \text{ to } Y_\alpha \text{ indexed by } \alpha \in A. \end{cases}$

The $\sigma$-algebra $\mathcal{M}(f.) \subseteq \mathcal{P}(X)$ given by $\mathcal{M}(f.) = \sigma\big(\{f_\alpha^{-1}(E) : E \in \mathcal{N}_\alpha \text{ and } \alpha \in A\}\big)$ is called the $\sigma$-algebra (on $X$) generated by $\{f_\alpha\}_{\alpha\in A}$. A *product $\sigma$-algebra* on $Y$ is the $\sigma$-algebra $\mathcal{M}(p.)$ generated by $\{p_\alpha\}_{\alpha\in A}$ when $X = Y$. We denote this $\sigma$-algebra $\mathcal{M}(p.)$ by $\bigotimes_{\alpha\in A} \mathcal{N}_\alpha$.

**Proposition 1.8.** *$f_\alpha$ is $(\mathcal{M}(f.), \mathcal{N}_\alpha)$-measurable for each $\alpha \in A$.*

*Proof.* Let $\mathcal{M}_\alpha \doteq \sigma(\{f_\alpha^{-1}(E) : E \in \mathcal{N}_a\})$. Then, Propositions 1.3 and 1.4 imply that $\mathcal{M}_\alpha$ is a $\sigma$-algebra and $f_\alpha$ is $(\mathcal{M}_\alpha, \mathcal{N}_\alpha)$-measurable. Since $\mathcal{M}_\alpha \subseteq \mathcal{M}(f.)$, the proof is completed by Corollary 1.1. □

**Corollary 1.3.** *$p_\alpha$ is $(\bigotimes_{\alpha\in A} \mathcal{N}_\alpha, \mathcal{N}_\alpha)$-measurable for each $\alpha \in A$.*

**Remark 1.3.** *Note that $\bigcap_{\alpha\in A} \mathcal{M}_\alpha$ is also a $\sigma$-algebra (on $X$) by Proposition 1.2, and $\bigcap_{\alpha\in A} \mathcal{M}_\alpha \subseteq \mathcal{M}(f.)$ holds by $\mathcal{M}_\alpha \subseteq \mathcal{M}(f.)$, where $\mathcal{M}_\alpha \doteq \sigma(\{f_\alpha^{-1}(E) : E \in \mathcal{N}_\alpha\})$. However, $\bigcap_{\alpha\in A} \mathcal{M}_\alpha = \mathcal{M}(f.)$ does not hold in general, and $f_\alpha$ is not necessarily $(\bigcap_{\alpha\in A} \mathcal{M}_\alpha, \mathcal{N}_\alpha)$-measurable.*

**Proposition 1.9.** *Suppose $\mathcal{E}_\alpha \subseteq \mathcal{P}(Y_\alpha)$ and $\mathcal{N}_\alpha = \sigma(\mathcal{E}_\alpha)$ for each $\alpha \in A$. Then,*

$$\bigotimes_{\alpha\in A} \mathcal{N}_\alpha = \sigma\Big(\Big\{p_\alpha^{-1}(E) : E \in \mathcal{E}_\alpha \text{ and } \alpha \in A\Big\}\Big). \tag{1.3}$$

*Moreover, if $Y_\alpha \in \mathcal{E}_\alpha$ for each $\alpha \in A$, then $\bigotimes_{\alpha\in A} \mathcal{N}_\alpha \subseteq \sigma\big(\big\{\prod_{\alpha\in A} E_\alpha : E_\alpha \in \mathcal{E}_\alpha\big\}\big)$. In addition to that, if $A$ is countable, then $\bigotimes_{\alpha\in A} \mathcal{N}_\alpha = \sigma\big(\big\{\prod_{\alpha\in A} E_\alpha : E_\alpha \in \mathcal{E}_\alpha\big\}\big)$.*

*Proof.* Let $\mathcal{F} \doteq \big\{p_\alpha^{-1}(E) : E \in \mathcal{E}_\alpha \text{ and } \alpha \in A\big\} \subseteq \mathcal{P}(Y)$. Then, obviously, $\mathcal{F} \subseteq \bigotimes_{\alpha\in A} \mathcal{N}_\alpha$, and thus we obtain $\sigma(\mathcal{F}) \subseteq \bigotimes_{\alpha\in A} \mathcal{N}_\alpha$ by Lemma 1.1. On the other hand, Lemma 1.3 implies that for each $\alpha \in A$, $\{E \subseteq Y_\alpha : p_\alpha^{-1}(E) \in \sigma(\mathcal{F})\}$ is a $\sigma$-algebra on $Y_\alpha$ that obviously contains $\mathcal{E}_\alpha$ (see the definition of $\mathcal{F}$) and thus, by Lemma 1.1, it also contains $\mathcal{N}_\alpha = \sigma(\mathcal{E}_\alpha)$. In other words, for each $\alpha \in A$, we have $p_\alpha^{-1}(E) \in \sigma(\mathcal{F})$ whenever $E \in \mathcal{N}_\alpha$, meaning that $\bigotimes_{\alpha\in A} \mathcal{N}_\alpha \subseteq \sigma(\mathcal{F})$ by Lemma 1.1. Therefore, $\bigotimes_{\alpha\in A} \mathcal{N}_\alpha = \sigma(\mathcal{F})$.

Next, let $\mathcal{G} \doteq \big\{\prod_{\alpha\in A} E_\alpha : E_\alpha \in \mathcal{E}_\alpha\big\}$ and $E \in \mathcal{E}_\alpha$. Then, since $p_\alpha^{-1}(E) = \prod_{\beta\in A} E_\beta$ where $E_\beta = Y_\beta \in \mathcal{E}_\beta$ for $\beta \neq \alpha$ and $E_\alpha = E \in \mathcal{E}_\alpha$, we have $\big\{p_\alpha^{-1}(E) : E \in \mathcal{E}_\alpha \text{ and } \alpha \in A\big\} \subseteq \mathcal{G} \subseteq \sigma(\mathcal{G})$ and thus $\bigotimes_{\alpha\in A} \mathcal{N}_\alpha \subseteq \sigma(\mathcal{G})$ by (1.3) and Lemma 1.1. On the other hand, let $E_\alpha \in \mathcal{E}_\alpha$ for each $\alpha \in A$ and suppose $A$ is countable.

Then, since any $\sigma$-algebra is closed under countable intersections (see Proposition 1.1), we have by (1.3) that $\bigcap_{\alpha\in A} p_\alpha^{-1}(E_\alpha) \in \bigotimes_{\alpha\in A} \mathcal{N}_\alpha$, where $\bigcap_{\alpha\in A} p_\alpha^{-1}(E_\alpha) = \prod_{\alpha\in A} E_\alpha$ holds and thus

$$\Big\{ \prod_{\alpha\in A} E_\alpha : E_\alpha \in \mathcal{E}_\alpha \Big\} \subseteq \bigotimes_{\alpha\in A} \mathcal{N}_\alpha.$$

Finally, by Lemma 1.1, $\sigma(\mathcal{G}) \subseteq \bigotimes_{\alpha\in A} \mathcal{N}_\alpha$ and therefore $\sigma(\mathcal{G}) = \bigotimes_{\alpha\in A} \mathcal{N}_\alpha$ when $A$ is countable. □

By letting $\mathcal{E}_\alpha = \mathcal{N}_\alpha$ for all $\alpha \in A$ in Proposition 1.9, we obtain the following corollary.

**Corollary 1.4.** $\bigotimes_{\alpha\in A} \mathcal{N}_\alpha \subseteq \sigma\big(\{ \prod_{\alpha\in A} N_\alpha : N_\alpha \in \mathcal{N}_\alpha\}\big)$. *Moreover, if $A$ is countable, then*

$$\bigotimes_{\alpha\in A} \mathcal{N}_\alpha = \sigma\bigg(\Big\{ \prod_{\alpha\in A} N_\alpha : N_\alpha \in \mathcal{N}_\alpha \Big\}\bigg).$$

*Proof.* For the completeness, we provide the independent proof, which is almost same to the second paragraph of the proof of Proposition 1.9. Let $\mathcal{N} \doteq \sigma\big(\{ \prod_{\alpha\in A} N_\alpha : N_\alpha \in \mathcal{N}_\alpha\}\big)$ and $N \in \mathcal{N}_\alpha$. Then, since $p_\alpha^{-1}(N) = \prod_{\beta\in A} N_\beta$ where $N_\beta = Y_\beta \in \mathcal{N}_\beta$ for $\beta \neq \alpha$ and $N_\alpha = N \in \mathcal{N}_\alpha$, we have $\big\{p_\alpha^{-1}(N) : N \in \mathcal{N}_\alpha \text{ and } \alpha \in A\big\} \subseteq \mathcal{N}$ and thus $\bigotimes_{\alpha\in A} \mathcal{N}_\alpha \subseteq \mathcal{N}$ by Lemma 1.1. On the other hand, let $N_\alpha \in \mathcal{N}_\alpha$ and suppose $A$ is countable. Then, since any $\sigma$-algebra is closed under countable intersections (see Proposition 1.1), $\bigcap_{\alpha\in A} p_\alpha^{-1}(N_\alpha) \in \bigotimes_{\alpha\in A} \mathcal{N}_\alpha$, where $\bigcap_{\alpha\in A} p_\alpha^{-1}(N_\alpha) = \prod_{\alpha\in A} N_\alpha$ holds and thus $\big\{ \prod_{\alpha\in A} N_\alpha : N_\alpha \in \mathcal{N}_\alpha\big\} \subseteq \bigotimes_{\alpha\in A} \mathcal{N}_\alpha$. Finally, by Lemma 1.1, we obtain $\mathcal{N} \subseteq \bigotimes_{\alpha\in A} \mathcal{N}_\alpha$ and therefore $\mathcal{N} = \bigotimes_{\alpha\in A} \mathcal{N}_\alpha$ when $A$ is countable. □

**Proposition 1.10.** *Let $(X, \mathcal{M})$ and $(Y_\alpha, \mathcal{N}_\alpha)$ $(\alpha \in A)$ be measurable spaces, and denote $\mathcal{N} \doteq \bigotimes_{\alpha\in A} \mathcal{N}_\alpha$. Then, $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable iff $f_\alpha = p_\alpha \circ f$ is $(\mathcal{M}, \mathcal{N}_\alpha)$-measurable for all $\alpha \in A$.*

*Proof.* Since every $p_\alpha$ is $(\mathcal{N}, \mathcal{N}_\alpha)$-measurable by Corollary 1.3 and the composition of measurable functions is also measurable by Proposition 1.7, $f_\alpha = p_\alpha \circ f : X \to Y_\alpha$ is $(\mathcal{M}, \mathcal{N}_\alpha)$-measurable for each $\alpha \in A$ if $f$ is $(\mathcal{M}, \mathcal{N})$-measurable. Conversely, if each $f_\alpha$ is $(\mathcal{M}, \mathcal{N}_\alpha)$-measurable, then for all $N \in \mathcal{N}_\alpha$ and each $\alpha \in A$, $f^{-1}(p_\alpha^{-1}(N)) = f_\alpha^{-1}(N) \in \mathcal{M}$; the proof is completed by $\mathcal{N} = \sigma\big(\{p_\alpha^{-1}(N) : N \in \mathcal{N}_\alpha \text{ and } \alpha \in A\}\big)$ and Proposition 1.6. □

### 1.3.1 The Special Case $A = \{1, 2\}$.

Now, consider the special case $A = \{1, 2\}$ with the two measurable spaces denoted $(X, \mathcal{M})$ and $(Y, \mathcal{N})$, which define the product $\sigma$-algebra $\mathcal{M} \otimes \mathcal{N}$ of $X \times Y$ and the rectangles $M \times N \subseteq X \times Y$ for $M \in \mathcal{M}$ and $N \in \mathcal{N}$. In this case, we obtain the following.

**Corollary 1.5.** *The collection $\mathcal{A}$ of finite disjoint unions of rectangles $M \times N$ with $M \in \mathcal{M}$ and $N \in \mathcal{N}$ is an algebra. Moreover, $\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{A}) = \sigma\big(\{M \times N : M \in \mathcal{M},\ N \in \mathcal{N}\}\big)$.*

*Proof.* Note that for any $M_1, M_2 \in \mathcal{M}$ and any $N_1, N_2 \in \mathcal{N}$,

$$(M_1 \times N_1) \cap (M_2 \times N_2) = \big\{(x, y) \in X \times Y : x \in M_1 \cap M_2 \text{ and } y \in N_1 \cap N_2\big\} = (M_1 \cap M_2) \times (N_1 \cap N_2),$$

and for any $M \in \mathcal{M}$ and $N \in \mathcal{N}$,

$$\begin{aligned}(M \times N)^c &= \{(x, y) \in X \times Y : x \notin M \text{ or } y \notin N\} \\ &= \big\{(x, y) \in X \times Y : (x \notin M \text{ and } y \in Y) \text{ or } (x \in X \text{ and } y \notin N)\big\} \\ &= (M^c \times Y) \cup (X \times N^c).\end{aligned}$$

Therefore, Proposition A.1 proves that the collection $\mathcal{A}$ of finite disjoint unions of rectangles $M \times N$ ($M \in \mathcal{M}$ and $N \in \mathcal{N}$) is an algebra. Moreover, let $\mathcal{R} \doteq \big\{M \times N : M \in \mathcal{M},\ N \in \mathcal{N}\big\}$. Then, $\mathcal{M} \otimes \mathcal{N} = \sigma\big(\mathcal{R}\big)$ by Corollary 1.4 and $\sigma(\mathcal{A}) = \sigma(\mathcal{R})$ is obvious as follows. $\sigma(\mathcal{R})$ is the $\sigma$-algebra that obviously contains any finite disjoint unions of rectangles in $\mathcal{R}$, implying $\mathcal{A} \subseteq \sigma(\mathcal{R})$ and by Lemma 1.1, $\sigma(\mathcal{A}) \subseteq \sigma(\mathcal{R})$. Conversely, it is obvious that $\mathcal{R} \subseteq \mathcal{A}$ and hence $\mathcal{R} \subseteq \sigma(\mathcal{A})$. Therefore, $\sigma(\mathcal{R}) \subseteq \sigma(\mathcal{A})$ again by Lemma 1.1. □

Now, for any subset $E \subseteq X \times Y$, define the $x$-section $E_x$ and $y$-section $E^y$ of $E$ by

$$E_x \doteq \big\{y \in Y : (x, y) \in E\big\} \text{ and } E^y \doteq \big\{x \in X : (x, y) \in E\big\}.$$

Similarly, for any function $f : X \times Y \to Z$ for a measurable space $(Z, \mathcal{O})$, define $f_x : Y \to Z$ for each $x \in X$ and $f^y : X \to Z$ for each $y \in Y$ as

$$f_x(y) = f^y(x) = f(x, y) \quad \forall (x, y) \in X \times Y.$$

**Proposition 1.11.** *If $E \in \mathcal{M} \otimes \mathcal{N}$, then $E_x \in \mathcal{N}$ for all $x \in X$ and $E^y \in \mathcal{M}$ for all $y \in Y$.*

*Proof.* Let $\overline{\mathcal{R}} \doteq \big\{E \subseteq X \times Y : E^y \in \mathcal{M} \text{ for all } y \in Y \text{ and } E_x \in \mathcal{N} \text{ for all } x \in X\big\}$. Then, $\overline{\mathcal{R}}$ contains $\mathcal{R}$ defined in the proof of Corollary 1.5 (e.g., $(M \times N)_x$ is equal to $N$ ($\in \mathcal{N}$) if $x \in M$ and $\varnothing$ ($\in \mathcal{N}$) if $x \notin M$). Moreover, for any $\{E_i\}_{i=1}^{\infty} \subseteq \overline{\mathcal{R}}$ and $E \in \overline{\mathcal{R}}$,

$$\begin{aligned}
\Big(\bigcup_{i=1}^{\infty} E_i\Big)_x &= \Big\{(x, y) \in X \times Y : (x, y) \in E_i \text{ for some } i \in \mathbb{N}\Big\}_x \\
&= \Big\{y \in Y : (x, y) \in E_i \text{ for some } i \in \mathbb{N}\Big\} = \bigcup_{i=1}^{\infty} (E_i)_x \in \mathcal{N} \\
(E^c)_x &= \Big\{(x, y) \in X \times Y : (x, y) \notin E\Big\}_x = \Big\{y \in Y : (x, y) \notin E\Big\} = (E_x)^c \in \mathcal{N}
\end{aligned}$$

and similarly, $\big(\bigcup_{i=1}^{\infty} E_i\big)^y = \bigcup_{i=1}^{\infty} (E_i)^y \in \mathcal{M}$ and $(E^c)^y = (E^y)^c \in \mathcal{M}$. Therefore, $\overline{\mathcal{R}}$ is a $\sigma$-algebra and thus by Lemma 1.1 and Corollary 1.5, $\overline{\mathcal{R}} \supseteq \sigma(\mathcal{R}) = \mathcal{M} \otimes \mathcal{N}$, which completes the proof. ☐

**Proposition 1.12.** *If $f$ is $(\mathcal{M} \otimes \mathcal{N}, \mathcal{O})$-measurable, then $f_x$ is $(\mathcal{N}, \mathcal{O})$-measurable for all $x \in X$ and $f^y$ is $(\mathcal{M}, \mathcal{O})$-measurable for all $y \in Y$.*

*Proof.* Let $E \doteq f^{-1}(F) \in \mathcal{M} \otimes \mathcal{N}$ for $F \in \mathcal{O}$. Then, for any $x \in X$, we have

$$E_x = \big\{f^{-1}(F)\big\}_x = \big\{(x, y) \in X \times Y : f(x, y) \in F\big\}_x = \big\{y \in Y : f(x, y) \in F\big\} = f_x^{-1}(F)$$

and thus by Proposition 1.11, $f_x^{-1}(F) = E_x \in \mathcal{N}$ for all $x \in X$, meaning that $f_x$ is $(\mathcal{N}, \mathcal{O})$-measurable for all $x \in X$. Similarly, for all $y \in Y$, $f^y$ is $(\mathcal{M}, \mathcal{O})$-measurable. ☐

## 1.4 Borel $\sigma$-Algebra, Some Topology, and Measurable Functions

The concept of the abstract Borel $\sigma$-algebra is essentially connected to the open sets and general topology. Thus, we begin this section with the general definition of topology and open/closed sets shown below.

**Definition 1.6.** *A topological space is an ordered pair $(Y, \mathcal{T})$, where $Y$ is a non-empty set and $\mathcal{T}$ is a collection of subsets of $Y$, satisfying the following axioms:*

*1)* $\varnothing, \; Y \in \mathcal{T}$,

*2)* $\{V_\alpha\}_{\alpha \in A} \subseteq \mathcal{T} \implies \bigcup_{\alpha \in A} V_\alpha \in \mathcal{T}$ *(closed under arbitrary unions),*

*3)* $V_1, V_2, \cdots, V_n \in \mathcal{T} \implies V_1 \cap V_2 \cap \cdots \cap V_n \in \mathcal{T}$ *(closed under finite intersections).*

*$\mathcal{T}$ is called a topology on $Y$; each member $V$ of $\mathcal{T}$ is said to be an open set; each complement $U^c$ is called a closed set. If the topology $\mathcal{T}$ is well-understood, we simply say that $Y$ is a topological space.*

**Definition 1.7.** *A Borel $\sigma$-algebra $\mathcal{B}$ on a topological space $Y$ is the $\sigma$-algebra generated by its topology $\mathcal{T}$ (the collection of all open sets), i.e., $\mathcal{B} \doteq \sigma(\mathcal{T})$.*

A countable intersection of open sets is called a $G_\delta$ set; a countable union of closed sets is called an $F_\sigma$ set; a countable union of $G_\delta$ sets is called a $G_{\delta\sigma}$ set; a countable intersection of $F_\sigma$ set is called an $F_{\sigma\delta}$ set (here, $\delta$ and $\sigma$ stand for intersection and union, respectively). By Proposition 1.1 and the definitions of a topology and a $\sigma$-algebra above, we obtain the following.

**Proposition 1.13.** *$\mathcal{B}$ contains the followings.*

1. *all of the open sets and closed sets;*
2. *arbitrary unions of open sets and arbitrary intersections of closed sets;*
3. *all of $G_\delta$ and $F_\sigma$ sets;*
4. *all of $G_{\delta\sigma}$ and $F_{\sigma\delta}$ sets.*

*Proof.* Since $\mathcal{T} \subseteq \sigma(\mathcal{T})$, $\mathcal{B}$ contains all of the open sets $U$ and their arbitrary unions. Since a $\sigma$-algebra is closed under taking complements, $\mathcal{B}$ also contains all of the closed sets $U^c$ and by De Morgan's laws, their arbitrary intersections. Moreover, a $\sigma$-algebra is closed under countable unions and intersections, and thereby $\mathcal{B}$ includes any $F_\sigma$ set (any countable union of closed sets) and $G_\delta$ set (any countable intersection of open sets). Likewise, any $F_{\sigma\delta}$ sets and $G_{\delta\sigma}$ sets also belong to $\mathcal{B}$. □

Such characterization of a Borel $\sigma$-algebra via topology gives the following property. In this section, whenever necessary, we denote $\mathcal{T}_X$ and $\mathcal{T}_Y$ the respective topologies on $X$ and $Y$; $\mathcal{B}_X \doteq \sigma(\mathcal{T}_X)$ and $\mathcal{B}_Y \doteq \sigma(\mathcal{T}_Y)$ their Borel $\sigma$-algebras; $U$ and $V$ open sets in $\mathcal{T}_X$ and $\mathcal{T}_Y$ (otherwise, $\mathcal{T}_Y$ and $\mathcal{B}_Y$ are abbreviated as $\mathcal{T}$ and $\mathcal{B}$), respectively.

**Proposition 1.14.** *Every continuous function $f : X \to Y$ is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable.*

*Proof.* $f$ is continuous iff $f^{-1}(V)$ is open in $X$ for every open set $V \in \mathcal{T}_Y$. Hence, for every $V \in \mathcal{T}_Y$, $f^{-1}(V) \in \mathcal{B}_X$ and the proof is complete by Propositoin 1.6. □

Similarly to Proposition 1.2, we obtain the following proposition for a family of topologies.

**Proposition 1.15.** *If $\{\mathcal{T}^\alpha\}_{\alpha\in A}$ is a family of topologies on $Y$, $\mathcal{T} = \bigcap_{\alpha\in A} \mathcal{T}^\alpha$ is also a topology on $Y$.*

*Proof.* Since $\varnothing$, $Y \in \mathcal{T}^\alpha$ for all $\alpha \in A$, they also belongs to $\mathcal{T}$. Next, suppose $\{V_\beta\}_{\beta\in B} \subseteq \mathcal{T}$, where $B$ is an index set. Then, $V_\beta \in \mathcal{T}$ implies that $V_\beta \in \mathcal{T}^\alpha$ for all $\beta \in B$. Since $\mathcal{T}^\alpha$ is a topology, we obtain $\bigcup_{\beta\in B} V_\beta \in \mathcal{T}^\alpha$ for all $\alpha \in A$, implying that $\bigcup_{\beta\in B} V_\beta \in \mathcal{T}$. Likewise, one can also show that $\mathcal{T}$ is closed under finite intersections. Therefore, $\mathcal{T}$ is a topology. □

**Definition 1.8.** *For any family $\mathcal{V} \subseteq \mathcal{P}(Y)$, $\mathcal{T}(\mathcal{V})$ denotes the smallest topology on $Y$ that contains $\mathcal{V}$; we call $\mathcal{T}(\mathcal{V})$ the topology generated by $\mathcal{V}$.*

**Remark 1.4.** *Note that for any family $\mathcal{V} \subseteq \mathcal{P}(Y)$, there is at least one topology, namely, the power set $\mathcal{P}(Y)$ itself, the largest topology on $Y$, that contains $\mathcal{V}$. Moreover, since any (uncountable) intersection of topologies is also a topology as shown in Proposition 1.15, the smallest topology $\mathcal{T}(\mathcal{V})$ in Definition 1.8 can be recognized as the intersection of all topologies containing $\mathcal{V}$. Moreover, if $\mathcal{V}$ is understood, one can construct such a topology $\mathcal{T}$ by the following procedure:*

1. *add $\varnothing$, every member of $\mathcal{V}$, and $Y$ to $\mathcal{T}$;*
2. *add all finite intersections of the sets in $\mathcal{T}$ to $\mathcal{T}$;*
3. *add all arbitrary unions of the sets in $\mathcal{T}$ to $\mathcal{T}$.*

*Here, the order of 1), 2) and 3) is strict and not interchangeable.*

**Proposition 1.16.** *Let $(Y, d)$ be a metric space and $B(y, r) \doteq \{z \in Y : d(x, z) < r\}$ be an open ball centerred at $y \in Y$ with its radius $r > 0$. Define $\mathcal{T}^d \subseteq \mathcal{P}(Y)$ as*

$$V \in \mathcal{T}^d \text{ iff } \forall y \in V : \; \exists r > 0 \text{ such that } B(y, r) \subseteq V. \tag{1.4}$$

*Then, $\mathcal{T}^d$ is a topology on $Y$. Moreover, every $V \in \mathcal{T}^d$ is a union of open balls.*

*Proof.* First, note that $Y \in \mathcal{T}^d$ is true since $B(y,r) \subseteq Y$ holds for any $r > 0$ by the definition of $B(y,r)$. Moreover, (1.4) is true vacuously (any statement that starts with $\forall x \in \varnothing$ is true, as there is no $x$ in the empty set to falsify the rest of the statement). Hence, $\varnothing \in \mathcal{T}^d$. Next, let $\{V_\alpha\}_{\alpha\in A} \subseteq \mathcal{T}^d$ and $y \in V_\beta$ for some $\beta \in A$. Then, there exists $r > 0$ such that $B(y,r) \subseteq V_\beta$ and hence, $B(y,r) \subseteq \bigcup_{\alpha\in A} V_\alpha$, that is, $\bigcup_{\alpha\in A} V_\alpha \in \mathcal{T}^d$. In a similar manner, suppose $V_1, V_2, \cdots, V_n \in \mathcal{T}^d$, with $\bigcap_{i=1}^n V_i \neq \varnothing$, and $y \in V_i$ for all $i = 1, 2, \cdots, n$. Then, for each $i$, there is $r_i > 0$ such that $B(y, r_i) \subseteq V_i$. Hence, $B(y,r) \subseteq V_i$ for all $i = 1, 2, \cdots, n$, where $r = \min\{r_i\}_{i=1}^n$. That is, $B(y,r) \subseteq \bigcap_{i=1}^n V_i$ and thereby $\bigcap_{i=1}^n V_i \in \mathcal{T}^d$. Therefore, $\mathcal{T}^d$ is a topology on $Y$.

Moreover, for any $V \in \mathcal{T}^d$, it is obvious that $\bigcup_{y\in V} B(y, r_y) \subseteq V$ for some $r_y > 0$ by (1.4). Conversely, since each $y \in V$ is obviously contained in $B(y, r_y)$, $V = \{y \in V\} \subseteq \bigcup_{y\in V} B(y, r_y)$. Therefore, every $V \in \mathcal{T}^d$ is the union of open balls $B(y, r_y)$. □

The topology $\mathcal{T}^d$ in Proposition 1.16 is called the topology on $Y$ induced by the metric $d$. Since every $V \in \mathcal{T}^d$ is a union of open balls by Proposition 1.16, the collection of all open balls

$$\mathcal{E} = \big\{B(y,r) : y \in Y \text{ and } r > 0\big\} \tag{1.5}$$

is a base of the topology $\mathcal{T}^d$.

**Definition 1.9.** *A collection $\mathcal{E}$ of subsets of $Y$ is said to be a base of a topology $\mathcal{T}$ on $Y$ iff every $V \in \mathcal{T}$ is a union of members of $\mathcal{E}$. The topological space $Y$ is said to be second countable iff its topology $\mathcal{T}$ has a countable base $\mathcal{E}$.*

**Definition 1.10.** *A subset $D$ of $Y$ is said to be dense iff every point $y \in Y$ either belongs to $D$ or is a limit point of $D$. A topological space $Y$ is separable if it contains a countable, dense subset $D$.*

**Proposition 1.17.** *A metric space is separable iff it is second countable.*

*Proof.* Every second countable space is also separable. To prove the necessity, suppose $Y$ is a separable metric space. Then, $Y$ contains a countable, dense subset $D$. The set $\mathbb{Q}$ of all rational numbers is countable, which implies

$$\mathfrak{B} \doteq \{B(y,r) : y \in D \text{ and } 0 < r \in \mathbb{Q}\}$$

is a countable collection of open balls. Let $\mathcal{T}^d$ be the topology on $Y$ induced by $d$ and $V \in \mathcal{T}^d$. Then, for each $\bar{y} \in V$, there is $\bar{r} > 0$ such that $B(\bar{y}, \bar{r}) \subseteq V$. Next, if $\bar{y} \in V \cap D$, set $y = \bar{y}$; otherwise, choose $y \in D$ such that $d(\bar{y}, y) < r$ for a rational number $r \in (0, \bar{r}/2]$. Then, by triangular inequality, for any $z \in B(y,r)$, we obtain $d(\bar{y}, z) \leq d(\bar{y}, y) + d(y, z) < 2r \leq \bar{r}$, which implies $z \in B(\bar{y}, \bar{r})$. In summary,

$$B(y,r) \subseteq B(\bar{y}, \bar{r}) \subseteq V \text{ and } \bar{y} \in B(y,r).$$

Therefore, for each $\bar{y} \in V$, there is $y \in D$ and a rational number $r > 0$ such that $\bar{y} \in B(y,r) \subseteq V$, where $B(y,r) \in \mathfrak{B}$. This implies

$$\bigcup_{\bar{y}\in V} B(y,r) \subseteq V = \{\bar{y} \in V\} \subseteq \bigcup_{\bar{y}\in V} B(y,r).$$

Therefore, $V = \bigcup_{\bar{y}\in V} B(y,r)$, that is, $Y$ is second countable. □

### 1.4.1 Borel σ-algebra on a Product Metric Space

Now, let $Y_i$ $(i = 1, 2, \cdots, n)$ be a metric space with its metric $d_i : Y_i \times Y_i \to \mathbb{R}_+$ and

$$Y \doteq \prod_{i=1}^n Y_i = Y_1 \times Y_2 \times \cdots \times Y_n.$$

Then, $Y$ is a metric space with the product metric $d : Y \times Y \to \mathbb{R}_+$:

$$d(y,z) \doteq \max\{d_1(y_1, z_1), d_2(y_2, z_2), \cdots, d_n(y_n, z_n)\}, \tag{1.6}$$

where $y = (y_1, y_2, \cdots, y_n) \in Y$ and $z = (z_1, z_2, \cdots, z_n) \in Y$. Let $\mathcal{B}_{Y_i} = \sigma(\mathcal{T}^{d_i})$ $(i = 1, 2, \cdots, n)$ and $\mathcal{B}_Y = \sigma(\mathcal{T}^d)$ be the Borel σ-algebras generated by the topologies $\mathcal{T}^{d_i}$ and $\mathcal{T}^d$ on $Y_i$ and $Y$ induced by the metrics $d_i$ and $d$, respectively. Denote the open balls in $Y$ and $Y_i$ $(i = 1, 2, \cdots, n)$ by $B(y,r)$ and $B_i(y_i, r)$ for $y \in Y$ and $y_i \in Y_i$, respectively.

**Lemma 1.4.** *$B(y,r) = \prod_{i=1}^n B_i(y_i, r)$ for each $y = (y_1, y_2, \cdots, y_n) \in Y$ and $r > 0$.*

*Proof.* Every $z = (z_1, z_2, \cdots, z_n) \in B(y,r)$ satisfies $d(y,z) < r$ which holds iff $d_i(y_i, z_i) < r$ for all $i = 1, 2, \cdots, n$ by the definition of the product metric (1.6). This directly proves $B(y,r) = \prod_{i=1}^n B_i(y_i, r)$. □

**Lemma 1.5.** *If $Y_i$'s are all separable, then so is $Y$.*[1]

*Proof.* For each $i = 1, 2, \cdots, n$, let $D_i \subseteq Y_i$ be a countable dense subset of $Y_i$ and $y = (y_1, y_2, \cdots, y_n) \in Y$ with each $y_i \in Y_i$. Then, since $D_i$ is dense in $Y_i$, there is a sequence $\{z_{ij}\}_{j=1}^\infty$ in $D_i$ such that

$$\lim_{j\to\infty} d_i(y_i, z_{ij}) = 0 \tag{1.7}$$

(if $y_i \in D_i$, $\{z_{ij}\}_{j=1}^\infty$ with $z_{ij} = y_i \in D_i$ for all $j$ is trivially such a sequence). For each $j \in \mathbb{N}$, let $z^{(j)} \doteq (z_{1j}, z_{2j}, \cdots, z_{nj}) \in D$, where $D \doteq \prod_{i=1}^n D_i$. Then, by (1.7) and the definition of the metric $d$,

$$\lim_{j\to\infty} d(y, z^{(j)}) = \lim_{j\to\infty} \Big( \max\{d_1(y_1, z_{1j}), d_2(y_2, z_{2j}), \cdots, d_n(y_n, z_{nj})\} \Big) = 0,$$

which implies that every $y \in Y$ is a limit point of $D$ and hence, $D$ is a dense subset of $Y$. Moreover, $D$ is countable since any finite product of countable subsets is also countable. Hence, $Y$ is separable. □

**Lemma 1.6.** *Every projection map $p_i : Y \to Y_i$ is continuous.*

*Proof.* Suppose $y,\ z \in Y$ and let $y_i,\ z_i \in Y_i$ be the $i$th elements of $y$ and $z$, respectively. Then, we have $p_i(y) = y_i$ and $p_i(z) = z_i$. Set $\delta = \varepsilon > 0$. Then, by the definition of the produce metric (1.6),

$$d(y,z) < \delta \implies d_i(p_i(y), p_i(z)) = d_i(y_i, z_i) < \varepsilon,$$

implying continuity of $p_i : (Y, d) \to (Y_i, d_i)$. □

**Proposition 1.18.** *$\bigotimes_{i=1}^n \mathcal{B}_{Y_i} \subseteq \mathcal{B}_Y$. Moreover, if $Y_i$'s are all separable, then $\bigotimes_{i=1}^n \mathcal{B}_{Y_i} = \mathcal{B}_Y$.*

*Proof.* By Proposition 1.9, we have $\bigotimes_{i=1}^n \mathcal{B}_{Y_i} = \sigma\big(\{\pi_i^{-1}(V_i) : V_i \in \mathcal{T}^{d_i} \text{ and } i = 1, 2, \cdots, n\}\big)$, where $V_i$ is open in $Y_i$. Since $p_i$ is continuous by Lemma 1.6, $\pi_i^{-1}(V_i)$ is open in $Y$. Hence, $\bigotimes_{i=1}^n \mathcal{B}_{Y_i} \subseteq \mathcal{B}_Y$ by Lemma 1.1. Moreover, suppose $Y_i$'s are all separable. Then, they are all second countable by Proposition 1.17, so that there exist countable bases $\mathfrak{B}_i \subseteq \mathcal{T}^{d_i}$, i.e., countable collection of open balls $B_i$, such that every $V_i \in \mathcal{T}^{d_i}$ is a union of members of $\mathfrak{B}_i$. Here, the union is actually a countable union since $\mathfrak{B}_i$ is countable, which yields $\mathcal{B}_{Y_i} = \sigma(\mathfrak{B}_i)$ $(i = 1, 2, \cdots, n)$. Since $Y$ is also second countable by Lemma 1.5 and Proposition 1.17, its base $\mathfrak{B}$ is a *countable* collection of open balls $B$. Since $B = \prod_{i=1}^n B_i$ by Lemma 1.4, we therefore obtain $\mathcal{B}_Y = \sigma\{\prod_{i=1}^n B_i : B_i \in \mathfrak{B}_i\}$ and hence by Proposition 1.9, $\bigotimes_{i=1}^n \mathcal{B}_{Y_i} = \mathcal{B}_Y$. □

Since $\mathbb{R}$ is separable, we obtain the following corollary for the Borel $\sigma$-algebra $\mathcal{B}_{\mathbb{R}^n} \doteq \sigma(\mathcal{T}_{\mathbb{R}^n})$, where $\mathcal{T}_{\mathbb{R}^n}$ is the standard topology on $\mathbb{R}^n$ induced by, e.g., the Euclidean metric $d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$.

**Corollary 1.6.** $\mathcal{B}_{\mathbb{R}^n} = \bigotimes_{i=1}^n \mathcal{B}_{\mathbb{R}}$.

### 1.4.2 Borel $\sigma$-algebra $\mathcal{B}_{\mathbb{R}}$ on $\mathbb{R}$

The Borel $\sigma$-algebra $\mathcal{B}_{\mathbb{R}}$ is the $\sigma$-algebra on $\mathbb{R}$ generated by the standard topology on $\mathbb{R}$, i.e., the topology induced by the distance metric $d(x,y) = |x-y|$ $(x, y \in \mathbb{R})$. As shown below, $\mathcal{B}_{\mathbb{R}}$ contains the sets of all open, closed, half-open intervals, the sets of all open and closed rays, and by Proposition 1.13, all of (arbitrary unions of) open sets, (arbitrary intersections of) closed sets, $G_\delta$, $F_\sigma$ sets, and $G_{\delta\sigma}$, $F_{\sigma\delta}$ sets on $\mathbb{R}$.

[1] In the most general situation, every countable product of separable topological spaces is also separable.

**Proposition 1.19.** *$\mathcal{B}_{\mathbb{R}} = \sigma(\mathcal{E}_i)$ for $\mathcal{E}_i$ $(i = 1, 2, 3, \cdots, 8)$ shown below:*

1. *the open intervals: $\mathcal{E}_1 = \{(a, b) : a < b)\}$,*
2. *the closed intervals: $\mathcal{E}_2 = \{[a, b] : a < b)\}$,*
3. *the half-open intervals: $\mathcal{E}_3 = \{(a, b] : a < b\}$, or $\mathcal{E}_4 = \{[a, b) : a < b\}$,*
4. *the open rays: $\mathcal{E}_5 = \{(a, \infty) : a \in \mathbb{R}\}$, or $\mathcal{E}_6 = \{(-\infty, b) : b \in \mathbb{R}\}$,*
5. *the closed rays: $\mathcal{E}_7 = \{[a, \infty) : a \in \mathbb{R}\}$, or $\mathcal{E}_8 = \{(-\infty, b] : b \in \mathbb{R}\}$.*

*Proof.* The elements $\mathcal{E}_i$ for $i \neq 3, 4$ are open or closed, and the elements of $\mathcal{E}_3$ and $\mathcal{E}_4$ are $G_\delta$-sets since $(a, b] = \bigcap_{n=1}^{\infty}(a, b + n^{-1})$ and $[a, b) = \bigcap_{n=1}^{\infty}(a - n^{-1}, b)$. That is, the elements of $\mathcal{E}_i$ all belong to $\mathcal{B}_{\mathbb{R}}$, so we have $\mathcal{E}_i \subset \mathcal{B}_{\mathbb{R}}$ and by Lemma 1.1, $\sigma(\mathcal{E}_i) \subseteq \mathcal{B}_{\mathbb{R}}$ for each $i = 1, 2, 3, \cdots, 8$. On the other hand, every open set in $\mathbb{R}$ is a countable union of open intervals $(a, b)$ and thereby, $\mathcal{T}_{\mathbb{R}} \subseteq \sigma(\mathcal{E}_1)$, where $\mathcal{T}_{\mathbb{R}}$ is the standard topology on $\mathbb{R}$, and by Lemma 1.1 again, $\mathcal{B}_{\mathbb{R}} \subseteq \sigma(\mathcal{E}_1)$, which proves $\mathcal{B}_{\mathbb{R}} = \sigma(\mathcal{E}_1)$. Moreover, for each $i \neq 1$, every open interval $(a, b) \in \mathcal{E}_1$ is a member of $\sigma(\mathcal{E}_i)$ as shown below:

$$(a, b) = \begin{cases} \bigcup_{n=1}^{\infty} [a + n^{-1}, b - n^{-1}] \in \sigma(\mathcal{E}_2), \\ \bigcup_{n=1}^{\infty} (a, b - n^{-1}] \in \sigma(\mathcal{E}_3) \text{ and } \bigcup_{n=1}^{\infty} [a + n^{-1}, b) \in \sigma(\mathcal{E}_4), \\ (a, \infty) \setminus [b, \infty) \in \sigma(\mathcal{E}_5) \text{ and } (-\infty, b) \setminus (-\infty, a] \in \sigma(\mathcal{E}_6), \\ \left(\bigcup_{n=1}^{\infty} [a + n^{-1}, \infty)\right) \setminus [b, \infty) \in \sigma(\mathcal{E}_7) \text{ and } \left(\bigcup_{n=1}^{\infty} (-\infty, b - n^{-1}]\right) \setminus (-\infty, a] \in \sigma(\mathcal{E}_8), \end{cases}$$

where we have applied Proposition 1.1 to show $(a, b) \in \sigma(\mathcal{E}_i)$ for each $i = 2, 3, 4, \cdots, 8$. Therefore, we have $\mathcal{E}_1 \subseteq \sigma(\mathcal{E}_i)$ and by Lemma 1.1, $\mathcal{B}_{\mathbb{R}} = \sigma(\mathcal{E}_1) \subseteq \sigma(\mathcal{E}_i)$; we conclude $\mathcal{B}_{\mathbb{R}} = \sigma(\mathcal{E}_i)$ for each $i = 2, 3, 4, \cdots, 8$. □

The following corollary is a result of applying Propositions 1.6 and 1.19; it characterizes an $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurable real-valued function $f : X \to \mathbb{R}$ by open/closed rays on $\mathbb{R}$.

**Corollary 1.7.** *The followings are equivalent for a function $f : X \to \mathbb{R}$ and a measurable space $(X, \mathcal{M})$.*

1. *$f$ is $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurable.*
2. *$f^{-1}((a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.*
3. *$f^{-1}([a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.*
4. *$f^{-1}((-\infty, b)) \in \mathcal{M}$ for all $b \in \mathbb{R}$.*
5. *$f^{-1}((-\infty, b]) \in \mathcal{M}$ for all $b \in \mathbb{R}$.*

**Proposition 1.20.** *$\mathcal{B}_{\mathbb{R}} \subset \mathcal{B}_{\overline{\mathbb{R}}}$.*

*Proof.* Note that $\mathcal{E}_5 \subset \overline{\mathcal{E}}_5 \subseteq \mathcal{B}_{\overline{\mathbb{R}}}$. Therefore, $\mathcal{B}_{\mathbb{R}} = \sigma(\mathcal{E}_5) \subseteq \mathcal{B}_{\overline{\mathbb{R}}}$ by Proposition 1.19 and Lemma 1.1, where the relation is strict since, for example, $\{\infty\} \in \mathcal{B}_{\overline{\mathbb{R}}}$ but $\{\infty\} \notin \mathcal{B}_{\mathbb{R}}$. □

### 1.4.3 Borel σ-algebra on the Extended Real Numbers

The Borel $\sigma$-algebra $\mathcal{B}_{\mathbb{R}}$ on $\mathbb{R}$ can be extended to that on the extended real numbers $\overline{\mathbb{R}} = [-\infty, \infty]$. Here, we adopt the conventions regarding $\pm\infty$:

$$\begin{array}{lll} x \pm \infty = \pm\infty \ (x \in \mathbb{R}), & -\infty - \infty = -\infty, & \infty + \infty = \infty, \\ x \cdot (\pm\infty) = \pm\infty \ (x > 0), & 0 \cdot (\pm\infty) = 0, & x \cdot (\pm\infty) = \mp\infty \ (x < 0) \end{array} \tag{1.8}$$

and define $\mathcal{B}_{\overline{\mathbb{R}}}$ as $\mathcal{B}_{\overline{\mathbb{R}}} \doteq \sigma(\mathcal{T}_{\overline{\mathbb{R}}})$, the $\sigma$-algebra on $\overline{\mathbb{R}}$ generated by the topology $\mathcal{T}_{\overline{\mathbb{R}}}$. Here, $\mathcal{T}_{\overline{\mathbb{R}}}$ is defined as the induced topology on $\overline{\mathbb{R}}$ by the metric $\bar{d} : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to \mathbb{R}$ given by

$$\bar{d}(x, y) \doteq \left| \tan^{-1} x - \tan^{-1} y \right| \text{ for } x, y \in \overline{\mathbb{R}}. \tag{1.9}$$

For notational convenience, we also denote $\mathcal{T}_\Pi$ the standard topology on $\Pi \doteq [-\pi/2, \pi/2]$, i.e., a topology on $\Pi$ induced by the distance metric $d(x, y) = |x - y|$.

**Lemma 1.7.** *The metric spaces* $(\overline{\mathbb{R}}, \bar{d})$ *and* $(\Pi, d)$ *are isometric.*

*Proof.* For each $x, y \in \Pi$, $d(x,y) = |x-y| = \left|\tan^{-1}(\tan x) - \tan^{-1}(\tan y)\right| = \bar{d}(\tan x, \tan y)$, where $\tan(\cdot)$, with the convention $\tan(\pm\pi/2) = \pm\infty$, is a one-to-one function from $\Pi$ to $\overline{\mathbb{R}}$. □

In the metric space $(\Pi, d)$, every open ball is expressed as $(a,b)$ $(-\pi/2 \le a < b \le \pi/2)$, $[-\pi/2, b)$ $(b \in \mathbb{R})$, or $(a, \pi/2]$ $(a \in \mathbb{R})$ and thus by Lemma 1.7, the open balls in $(\overline{\mathbb{R}}, \bar{d})$ are of the forms $(a,b)$, $[-\infty, b)$, and $(a, \infty]$. For notational convenience, we define $\mathcal{E}_{\overline{\mathbb{R}}}$ as

$$\mathcal{E}_{\overline{\mathbb{R}}} \doteq \{(a,b) : -\infty \le a < b \le \infty\} \cup \{(a,\infty],\, [-\infty, b) : a, b \in \mathbb{R}\} = \mathcal{E}_1 \cup \mathcal{E}_5 \cup \mathcal{E}_6 \cup \overline{\mathcal{E}}_5 \cup \overline{\mathcal{E}}_6,$$

where $\mathcal{E}_1$, $\mathcal{E}_5$, and $\mathcal{E}_6$ are open intervals/rays defined on $\mathbb{R}$ in Proposition 1.19, and the extended open rays $\overline{\mathcal{E}}_5$, and $\overline{\mathcal{E}}_6$ are given by

$$\overline{\mathcal{E}}_5 \doteq \big\{(a,\infty] : a \in \mathbb{R}\big\}, \text{ and } \overline{\mathcal{E}}_6 \doteq \big\{[-\infty, b) : b \in \mathbb{R}\big\}.$$

Then, by combining this with Proposition 1.16, we obtain the following corollary.

**Corollary 1.8.** $\mathcal{O} \in \mathcal{T}_{\overline{\mathbb{R}}}$ *iff for all* $x \in \mathcal{O}$*, there exists* $I_x \in \mathcal{E}_{\overline{\mathbb{R}}}$ *such that* $I_x \subseteq \mathcal{O}$*. Moreover, every* $\mathcal{O} \in \mathcal{T}_{\overline{\mathbb{R}}}$ *is a union of members of* $\mathcal{E}_{\overline{\mathbb{R}}}$*.*

Denote $\overline{\mathcal{E}}_{\overline{\mathbb{R}}}$ the set of all closed intervals in $\overline{\mathbb{R}}$, i.e., $\overline{\mathcal{E}}_{\overline{\mathbb{R}}} \doteq \big\{[a,b] : -\infty \le a \le b \le \infty\big\}$. Then, by Proposition 1.13, and Corollary 1.8, $\mathcal{B}_{\overline{\mathbb{R}}}$ contains arbitrary unions of members of $\mathcal{E}_{\overline{\mathbb{R}}}$ (i.e., open sets), arbitrary intersections of members of $\overline{\mathcal{E}}_{\overline{\mathbb{R}}}$ (i.e., closed sets), and all of their corresponding $G_\delta$, $F_\sigma$, $G_{\delta\sigma}$ and $F_{\sigma\delta}$ sets. It also contains every member of $\mathcal{B}_{\mathbb{R}}$. Moreover, the next proposition states the useful form of $\mathcal{B}_{\overline{\mathbb{R}}}$ expressed in terms of $\mathcal{B}_{\mathbb{R}}$.

**Proposition 1.21.** $\mathcal{B}_{\overline{\mathbb{R}}} = \big\{E \subseteq \overline{\mathbb{R}} : E \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}\big\}$.

*Proof.* Let $\overline{\mathcal{B}} \doteq \big\{E \subseteq \overline{\mathbb{R}} : E \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}\big\}$. Then, $\overline{\mathcal{B}}$ is a $\sigma$-algebra, since so is $\mathcal{B}_{\mathbb{R}}$. Obviously, for $\{E_i\}_{i=1}^\infty \subset \mathcal{B}_{\overline{\mathbb{R}}}$, we have $E_i \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}$ for each $i \in \mathbb{N}$ and thereby, $\left(\bigcup_{i=1}^\infty E_i\right) \cap \mathbb{R} = \bigcup_{i=1}^\infty (E_i \cap \mathbb{R}) \in \mathcal{B}_{\mathbb{R}}$, implying $\bigcup_{i=1}^\infty E_i \in \overline{\mathcal{B}}$; for $E \in \mathcal{B}_{\overline{\mathbb{R}}}$, we have $E \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}$ and thus $\mathbb{R} \cap E^c = \mathbb{R} \cap (E^c \cup \{-\infty, \infty\}) = \mathbb{R} \setminus (E \cap \mathbb{R}) \in \mathcal{B}_{\mathbb{R}}$ by $\mathbb{R} \in \mathcal{B}_{\mathbb{R}}$ and Proposition 1.1. Hence, $\overline{\mathcal{B}}$ is a $\sigma$-algebra.

Now, suppose $E \in \overline{\mathcal{B}}$. Then, since $\{\pm\infty\}$, $\{-\infty, \infty\} \in \mathcal{B}_{\overline{\mathbb{R}}}$ ($\because$ $\{\pm\infty\} \in \overline{\mathcal{E}}_{\overline{\mathbb{R}}}$), we have

$$E = E \cap \overline{\mathbb{R}} = \underbrace{(E \cap \{-\infty, \infty\})}_{=\{\pm\infty\} \text{ or } \{-\infty,\infty\}} \cup \underbrace{(E \cap \mathbb{R})}_{\in \mathcal{B}_{\mathbb{R}} \subset \mathcal{B}_{\overline{\mathbb{R}}}} \in \mathcal{B}_{\overline{\mathbb{R}}},$$

implying $\overline{\mathcal{B}} \subseteq \mathcal{B}_{\overline{\mathbb{R}}}$. To prove the converse, let $E \in \mathcal{T}_{\overline{\mathbb{R}}}$. Then, by Corollary 1.8, there exists $\big\{I_x \in \mathcal{E}_{\overline{\mathbb{R}}}\big\}_{x \in E}$ such that $E = \bigcup_{x \in E} I_x$ and thus, $E \cap \mathbb{R} = \bigcup_{x \in E} I_x \cap \mathbb{R}$, where each $I_x \cap \mathbb{R}$ is an open ball in $\mathbb{R}$. Hence, $E \cap \mathbb{R} \in \mathcal{T}_{\mathbb{R}}$ and hence $E \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}$ $(= \sigma(\mathcal{T}_{\mathbb{R}}))$. That is, $\mathcal{T}_{\overline{\mathbb{R}}} \subseteq \overline{\mathcal{B}}$. This implies $\mathcal{B}_{\overline{\mathbb{R}}} \subseteq \overline{\mathcal{B}}$ by Lemma 1.1, hence we conclude $\mathcal{B}_{\overline{\mathbb{R}}} = \overline{\mathcal{B}}$. □

**Proposition 1.22.** $\mathcal{B}_{\mathbb{R}} = \{E \cap \mathbb{R} : E \in \mathcal{B}_{\overline{\mathbb{R}}}\}$.

*Proof.* Let $\mathcal{B} \doteq \{E \cap \mathbb{R} : E \in \mathcal{B}_{\overline{\mathbb{R}}}\}$. Then, since we have $E \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}$ for each $E \in \mathcal{B}_{\overline{\mathbb{R}}}$ by Proposition 1.21, we have $\mathcal{B} \subseteq \mathcal{B}_{\mathbb{R}}$. Moreover, $\mathcal{B}_{\mathbb{R}} = \{E \cap \mathbb{R} : E \in \mathcal{B}_{\mathbb{R}}\}$ trivially holds ($\because$ $E \cap \mathbb{R} = E$ for any $E \in \mathcal{B}_{\mathbb{R}}$), hence by $\mathcal{B}_{\mathbb{R}} \subset \mathcal{B}_{\overline{\mathbb{R}}}$ (Proposition 1.20), we have $\mathcal{B}_{\mathbb{R}} \subseteq \mathcal{B}$. Therefore, we conclude $\mathcal{B}_{\mathbb{R}} = \mathcal{B}$. □

The next proposition regarding $\mathcal{B}_{\overline{\mathbb{R}}}$ is parallel to Proposition 1.19 regarding $\mathcal{B}_{\mathbb{R}}$.

**Proposition 1.23.** $\mathcal{B}_{\overline{\mathbb{R}}} = \sigma(\mathcal{E}_{\overline{\mathbb{R}}}) = \sigma(\overline{\mathcal{E}}_i)$ *for* $i = 5, 6$.

*Proof.* Every member of $\mathcal{E}_{\overline{\mathbb{R}}}$ and $\overline{\mathcal{E}}_i$ for $i = 5, 6$ is an open set, i.e., belongs to the topology $\mathcal{T}_{\overline{\mathbb{R}}}$. Hence, since $\mathcal{B}_{\overline{\mathbb{R}}} = \sigma(\mathcal{T}_{\overline{\mathbb{R}}})$, we have $\mathcal{E}_{\overline{\mathbb{R}}} \subseteq \mathcal{B}_{\overline{\mathbb{R}}}$ and $\overline{\mathcal{E}}_i \subseteq \mathcal{B}_{\overline{\mathbb{R}}}$ and by Lemma 1.1, $\sigma(\mathcal{E}_{\overline{\mathbb{R}}}) \subseteq \mathcal{B}_{\overline{\mathbb{R}}}$ and $\sigma(\overline{\mathcal{E}}_i) \subseteq \mathcal{B}_{\overline{\mathbb{R}}}$ for $i = 5, 6$, respectively. On the other hand, every $\mathcal{O} \in \mathcal{T}_{\overline{\mathbb{R}}}$ is a countable union of members of $\mathcal{E}_{\overline{\mathbb{R}}}$ by Corollary 1.8

and thereby, $\mathcal{T}_{\overline{\mathbb{R}}} \subseteq \sigma(\mathcal{E}_{\overline{\mathbb{R}}})$, and by Lemma 1.1 again, $\mathcal{B}_{\overline{\mathbb{R}}} \subseteq \sigma(\mathcal{E}_{\overline{\mathbb{R}}})$, which proves $\mathcal{B}_{\overline{\mathbb{R}}} = \sigma(\mathcal{E}_{\overline{\mathbb{R}}})$. Moreover, for $i = 5, 6$, every member of $\mathcal{E}_{\overline{\mathbb{R}}}$ is a member of $\sigma(\mathcal{E}_j)$ as shown below:

$$\begin{aligned}
[-\infty, b) &= \textstyle\bigcup_{n\in\mathbb{N}} [-\infty, b - n^{-1}] = \bigcup_{n\in\mathbb{N}} (b - n^{-1}, \infty]^c \in \sigma(\mathcal{E}_5) \\
(a, \infty] &= \textstyle\bigcup_{n\in\mathbb{N}} [a + n^{-1}, \infty] = \bigcup_{n\in\mathbb{N}} [-\infty, a + n^{-1})^c \in \sigma(\mathcal{E}_6) \\
(a, b) = [-\infty, b) \cap (a, \infty] &= \begin{cases} \left(\bigcup_{n\in\mathbb{N}} (b - n^{-1}, \infty]^c\right) \cap (a, \infty] \in \sigma(\mathcal{E}_5), \\ [-\infty, b) \cap \left(\bigcup_{n\in\mathbb{N}} [-\infty, a + n^{-1})^c\right) \in \sigma(\mathcal{E}_6), \end{cases}
\end{aligned}$$

where we have applied Proposition 1.1. Therefore, we have $\mathcal{E}_{\overline{\mathbb{R}}} \subseteq \sigma(\mathcal{E}_i)$ and thus $\mathcal{B}_{\overline{\mathbb{R}}} = \sigma(\mathcal{E}_{\overline{\mathbb{R}}}) \subseteq \sigma(\mathcal{E}_i)$ by Lemma 1.1; we conclude $\mathcal{B}_{\overline{\mathbb{R}}} = \sigma(\mathcal{E}_i)$ for each $i = 5, 6$. □

## 1.5 Extended-real-valued Measurable Functions

Now, we consider extended-real-valued functions $f : X \to \overline{\mathbb{R}}$ and the measurable spaces $(X, \mathcal{M})$ and $(\overline{\mathbb{R}}, \mathcal{B}_{\overline{\mathbb{R}}})$. From now on, we say that an extended-real-valued function $f$ is $\mathcal{M}$-measurable iff it is $(\mathcal{M}, \mathcal{B}_{\overline{\mathbb{R}}})$-measurable, or just measurable if $\mathcal{M}$ is understood so that there is no confusion.

The following corollary is parallel to Corollary 1.7 and is a result of applying Propositions 1.6 and 1.23.

**Corollary 1.9.** *$f$ is $\mathcal{M}$-measurable iff $f^{-1}((a, \infty]) \in \mathcal{M}$ for all $a \in \mathbb{R}$ iff $f^{-1}([-\infty, b)) \in \mathcal{M}$ for all $b \in \mathbb{R}$.*

**Lemma 1.8.** *Let $f : X \to \overline{\mathbb{R}}$ and $Y = f^{-1}(\mathbb{R})$. Then, $f$ is $\mathcal{M}$-measurable iff*

1. *$f^{-1}(\{-\infty\}) \in \mathcal{M}$ and $f^{-1}(\{\infty\}) \in \mathcal{M}$;*
2. *$f$ is $\mathcal{M}$-measurable on $Y$, i.e., $f^{-1}(E) \cap Y \in \mathcal{M}$ for every $E \in \mathcal{B}_{\overline{\mathbb{R}}}$.*

*Proof.* ($\Longrightarrow$) Suppose $f$ is $\mathcal{M}$-measurable. Then, since $\mathbb{R} \in \mathcal{B}_{\overline{\mathbb{R}}}$ by Proposition 1.21, we have $f^{-1}(\mathbb{R}) \in \mathcal{M}$. Therefore, for any $E \in \mathcal{B}_{\overline{\mathbb{R}}}$, we have $f^{-1}(E) \cap Y \in \mathcal{M}$ by Proposition 1.1. That is, $f$ is $\mathcal{M}$-measurable on $Y$. Moreover, we have $f^{-1}(\{-\infty\}) \in \mathcal{M}$ and $f^{-1}(\{\infty\}) \in \mathcal{M}$ as follows:

$$\begin{aligned}
f^{-1}(\{-\infty\}) &= f^{-1}\Big(\textstyle\bigcap_{n=1}^{\infty} [-\infty, -n)\Big) = \bigcap_{n=1}^{\infty} f^{-1}\big([-\infty, -n)\big) \in \mathcal{M}, \\
f^{-1}(\{\infty\}) &= f^{-1}\Big(\textstyle\bigcap_{n=1}^{\infty} (n, \infty]\Big) = \bigcap_{n=1}^{\infty} f^{-1}\big((n, \infty]\big) \in \mathcal{M},
\end{aligned}$$

where we have applied Lemma 1.2, Corollary 1.9 and Proposition 1.1.

($\Longleftarrow$) Suppose $f$ is $\mathcal{M}$-measurable on $Y$, $f^{-1}(\{-\infty\}) \in \mathcal{M}$, and $f^{-1}(\{\infty\}) \in \mathcal{M}$. We know that

$$f^{-1}\big((a, \infty]\big) = f^{-1}\Big(\big((a, \infty] \cap \mathbb{R}\big) \cup \{\infty\}\Big) = \Big(f^{-1}\big((a, \infty]\big) \cap Y\Big) \cup f^{-1}(\{\infty\})$$

for any $a \in \mathbb{R}$ by Lemma 1.2. Here, since $f$ is measurable on $Y$, we have $f^{-1}\big((a, \infty]\big) \cap Y \in \mathcal{M}$ and hence $f^{-1}\big((a, \infty]\big) \in \mathcal{M}$ for any $a \in \mathbb{R}$. Therefore, $f$ is $\mathcal{M}$-measurable by Corollary 1.9. □

**Proposition 1.24.** *$f$ is $\mathcal{M}$-measurable iff $f^{-1}(\{\infty\}) \in \mathcal{M}$, $f^{-1}(\{-\infty\}) \in \mathcal{M}$, and*

$$f^{-1}(E) \in \mathcal{M} \text{ for each } E \in \mathcal{B}_{\mathbb{R}}.$$

*Proof.* Let $Y \doteq f^{-1}(\mathbb{R})$. Then, if $f^{-1}(E) \cap Y \in \mathcal{M}$ for all $E \in \mathcal{B}_{\overline{\mathbb{R}}}$, then we have by Lemma 1.2

$$f^{-1}(E \cap \mathbb{R}) \in \mathcal{M} \quad \text{for all } E \in \mathcal{B}_{\overline{\mathbb{R}}}, \tag{1.10}$$

which implies $f^{-1}(F) \in \mathcal{M}$ for all $F \in \mathcal{B}_{\mathbb{R}}$ by Proposition 1.22. Conversely, if $f^{-1}(F) \in \mathcal{M}$ for all $F \in \mathcal{B}_{\mathbb{R}}$, then $f^{-1}(E \cap \mathbb{R}) \in \mathcal{M}$ for all $E \in \mathcal{B}_{\overline{\mathbb{R}}}$ by Proposition 1.21 (or Proposition 1.22), which implies $f^{-1}(E) \cap Y \in \mathcal{M}$ for all $E \in \mathcal{B}_{\overline{\mathbb{R}}}$ by Lemma 1.2. Therefore, the proof is completed by applying Lemma 1.8 above. □

In what follows, we provide the properties of $\overline{\mathbb{R}}$-valued measurable functions.

**Lemma 1.9.** *If $X = A \cup B$ where $A, B \in \mathcal{M}$, then a function $f : X \to \overline{\mathbb{R}}$ is $\mathcal{M}$-measurable iff $f$ is measurable on $A$ and $B$.*

*Proof.* Suppose $f$ is $\mathcal{M}$-measurable. Then, $f^{-1}(E) \in \mathcal{M}$ for each $E \in \mathcal{B}_{\overline{\mathbb{R}}}$ and hence, $f^{-1}(E) \cap A \in \mathcal{M}$ and $f^{-1}(E) \cap B \in \mathcal{M}$ by Proposition 1.1. Conversely, if $f$ is $\mathcal{M}$-measurable on $A$ and $B$, we have

$$f^{-1}(E) = f^{-1}(E) \cap (A \cup B) = (f^{-1}(E) \cap A) \cup (f^{-1}(E) \cap B) \in \mathcal{M},$$

for each $E \in \mathcal{B}_{\overline{\mathbb{R}}}$, hence $f$ is $\mathcal{M}$-measurable. □

**Proposition 1.25.** *Suppose $f, g : X \to \overline{\mathbb{R}}$ are $\mathcal{M}$-measurable. Then,*

1. *$fg$ is $\mathcal{M}$-measurable (with $0 \cdot (\pm\infty) = 0$).*
2. *If there does not exist $x \in X$ such that $f(x) = -g(x) = \pm\infty$, then $f + g$ is $\mathcal{M}$-measurable. Generally, if $h$ is defined for some $a \in \overline{\mathbb{R}}$ as $h(x) = a$ if $f(x) = -g(x) = \pm\infty$ and $h(x) = f(x) + g(x)$ otherwise, then $h$ is $\mathcal{M}$-measurable.*

*Proof.* Let $F : X \to \overline{\mathbb{R}}^2$, $\phi, \psi : \overline{\mathbb{R}}^2 \to \overline{\mathbb{R}}$ be defined as $F(x) \doteq (f(x), g(x))$, $\phi(u, v) \doteq u + v$ and $\psi(u, v) \doteq uv$, with the convention $0 \cdot (\pm\infty) = 0$. Since $\overline{\mathbb{R}}$ is separable, we have $\mathcal{B}_{\overline{\mathbb{R}}^2} = \mathcal{B}_{\overline{\mathbb{R}}} \otimes \mathcal{B}_{\overline{\mathbb{R}}}$ by Proposition 1.18. Hence, $F$ is $(\mathcal{M}, \mathcal{B}_{\overline{\mathbb{R}}^2})$-measurable by Proposition 1.10. Moreover, by continuity and Proposition 1.14, $\psi$ is $(\mathcal{B}_{\overline{\mathbb{R}}^2}, \mathcal{B}_{\overline{\mathbb{R}}})$-measurable. For $\phi$, let $A = \{(-\infty, \infty), (\infty, -\infty)\} \subset \overline{\mathbb{R}}^2$ and $B = X \setminus A$. Then, by Proposition 1.1, since

$$(-\infty, \infty) = \bigcap_{n,m \in \mathbb{N}} [-\infty, n) \times (m, \infty] \in \mathcal{B}_{\overline{\mathbb{R}}^2} \text{ and } (\infty, -\infty) = \bigcap_{n,m \in \mathbb{N}} (n, \infty] \times [-\infty, m) \in \mathcal{B}_{\overline{\mathbb{R}}^2},$$

we have $A = \{(-\infty, \infty)\} \cup \{(\infty, -\infty)\} \in \mathcal{B}_{\overline{\mathbb{R}}^2}$ and $B \in \mathcal{B}_{\overline{\mathbb{R}}^2}$. Hence, for any $E \in \mathcal{B}_{\overline{\mathbb{R}}}$, $\phi^{-1}(E) \cap A$ is equal to $\emptyset$, $\{(\infty, -\infty)\}$, $\{(-\infty, \infty)\}$, or $A$, all of which belong to $\mathcal{B}_{\overline{\mathbb{R}}^2}$. Moreover, by continuity of $\phi$ over $B$ and Proposition 1.14,

$$\phi^{-1}(E) \cap B = \phi|_B^{-1}(E) \in \mathcal{B}_{\overline{\mathbb{R}}^2}$$

for each $E \in \mathcal{B}_{\overline{\mathbb{R}}}$. Therefore, $\phi$ is $\mathcal{B}_{\overline{\mathbb{R}}^2}$-measurable by Lemma 1.9. By the measurability of $\psi$, $\phi$, and $F$, we conclude that $fg = \psi \circ F$ and $h = \phi \circ F$ are $\mathcal{M}$-measurable. □

**Proposition 1.26.** *Let $\{f_i : X \to \overline{\mathbb{R}}\}_{i=1}^{\infty}$ be a sequence of $\overline{\mathbb{R}}$-valued $\mathcal{M}$-measurable functions. Then,*

$$g_1(x) = \sup_{i \in \mathbb{N}} f_i(x) \qquad g_3(x) = \limsup_{i \to \infty} f_i(x)$$
$$g_2(x) = \inf_{i \in \mathbb{N}} f_i(x) \qquad g_4(x) = \liminf_{i \to \infty} f_i(x)$$

*are all $\mathcal{M}$-measurable. If $f(x) = \lim_{i \to \infty} f_i(x)$ exists for every $x \in X$, then $f$ is $\mathcal{M}$-measurable.*

*Proof.* For any $a \in \mathbb{R}$, we have

$$g_1^{-1}((a, \infty]) = \Big\{x : \sup_{i \in \mathbb{N}} f_i(x) > a\Big\} = \Big\{x : f_i(x) > a \text{ for some } i \in \mathbb{N}\Big\} = \bigcup_{i=1}^{\infty} \Big\{x : f_i(x) > a\Big\} = \bigcup_{i=1}^{\infty} f_i^{-1}((a, \infty])$$
$$g_2^{-1}([-\infty, a)) = \Big\{x : \inf_{i \in \mathbb{N}} f_i(x) < a\Big\} = \Big\{x : f_i(x) < a \text{ for some } i \in \mathbb{N}\Big\} = \bigcup_{i=1}^{\infty} \Big\{x : f_i(x) < a\Big\} = \bigcup_{i=1}^{\infty} f_i^{-1}([-\infty, a));$$

As $f_i$ is $\mathcal{M}$-measurable, $f_i^{-1}((a, \infty])$, $f_i^{-1}([-\infty, a)) \in \mathcal{M}$ for all $i \in \mathbb{N}$, hence we have $g_1^{-1}((a, \infty]) \in \mathcal{M}$ and $g_2^{-1}([-\infty, a)) \in \mathcal{M}$ for each $a \in \mathbb{R}$. Therefore, by Corollary 1.9, $g_1$ and $g_2$ are $\mathcal{M}$-measurable. Likewise, $h_j$ defined as $h_j(x) \doteq \sup_{i > j} f_i(x)$ is $\mathcal{M}$-measurable and hence, $g_3 = \inf_{j \in \mathbb{N}} \big(\sup_{i > j} f_i\big) = \inf_{j \in \mathbb{N}} h_j$ is $\mathcal{M}$-measurable. One can also prove in a similar way that $g_4$ is $\mathcal{M}$-measurable. Finally, if $f$ exists, then $f = g_3 = g_4$, so $f$ is $\mathcal{M}$-measurable. □

**Corollary 1.10.** *If $f, g : X \to \overline{\mathbb{R}}$ are $\mathcal{M}$-measurable, then so are $\max(f, g)$ and $\min(f, g)$.*

In this textbook, we identify a function $f : X \to Y$ for $Y \subseteq \overline{\mathbb{R}}$ as

$$\text{a function } f : X \to \overline{\mathbb{R}} \text{ whose image is a subset of } Y \text{ (i.e., } \mathrm{Im}(f) \subseteq Y). \tag{1.11}$$

This clarification defines the inverse map $f^{-1} : \mathcal{P}(\overline{\mathbb{R}}) \to \mathcal{P}(X)$, with the property that $f^{-1}(E) = \varnothing$ for any $E \in \mathcal{P}(Y^c)$. This property on $Y^c$ ensures that checking the measurability on $f^{-1}(Y)$ is enough for measurability on $X$ as shown below.

**Proposition 1.27.** *A function $f : X \to Y$ for $Y \in \mathcal{B}_{\overline{\mathbb{R}}}$ is $\mathcal{M}$-measurable iff it is $\mathcal{M}$-measurable on $f^{-1}(Y)$, i.e., iff $f^{-1}(E \cap Y) \in \mathcal{M}$ for all $E \in \mathcal{B}_{\overline{\mathbb{R}}}$.*

*Proof.* Apply Lemma 1.9 with $A = f^{-1}(Y) \in \mathcal{M}$ and $B = f^{-1}(Y^c) = \varnothing \in \mathcal{M}$; note that $f^{-1}(E) \cap A = f^{-1}(E \cap Y) \in \mathcal{M}$ for all $E \subseteq \overline{\mathbb{R}}$. □

Now, any $\mathcal{M}$-measurable function $f : X \to \overline{\mathbb{R}}$ can be decomposed as two extended positive real-valued functions $f^+, f^- : X \to Y$ with $Y = [0, \infty]$ as

$$f^+(x) \doteq \max(f(x), 0), \quad f^-(x) \doteq \max(-f(x), 0) \tag{1.12}$$

which are $\mathcal{M}$-measurable by Corollary 1.10 and satisfy $f = f^+ - f^-$. Since at least one of $f^+(x)$ and $f^-(x)$ is finite for each $x \in X$, $f$ is $\mathcal{M}$-measurable iff so are $f^+$ and $f^-$ by Proposition 1.25. Throughout the textbook, we denote

$$\mathcal{L}^+ \equiv \mathcal{L}^+(X, \mathcal{M}) \doteq \text{the space of all } \mathcal{M}\text{-measurable functions } f : X \to [0, \infty]. \tag{1.13}$$

Then, the above statement can be rephrased as follows.

**Corollary 1.11.** *$f : X \to \overline{\mathbb{R}}$ is $\mathcal{M}$-measurable iff $f^+, f^- \in \mathcal{L}^+$.*

### 1.5.1 Approximation by Simple Functions

Now, we show that for any measurable function $f : X \to \overline{\mathbb{R}}$ on a measurable space $(X, \mathcal{M})$, there exists a monotonic sequence of simple functions $\{\phi_n\}$ which pointwisely (and uniformly when $f$ is bounded) converges to $f$. To define a simple function, we introduce the indicator function $\mathbf{1}_E : X \to \{0, 1\}$ for a subset $E \subseteq X$:

$$\mathbf{1}_E(x) \doteq \begin{cases} 1 \text{ if } x \in E, \\ 0 \text{ if } x \notin E. \end{cases}$$

**Lemma 1.10.** *$\mathbf{1}_E$ is $\mathcal{M}$-measurable iff $E \in \mathcal{M}$.*

*Proof.* If $E \in \mathcal{M}$, then we have $\mathbf{1}_E^{-1}\{1\} = E$, $\mathbf{1}_E^{-1}\{0\} = E^c$, $\mathbf{1}_E^{-1}\{0, 1\} = X$, and $\mathbf{1}_E^{-1}(\varnothing) = \varnothing$, all of which belong to the $\sigma$-algebra $\mathcal{M}$. Conversely, if $\mathbf{1}_E$ is $\mathcal{M}$-measurable, then $E = \mathbf{1}_E^{-1}\{1\} \in \mathcal{M}$. □

**Proposition 1.28.** *If $f : X \to \overline{\mathbb{R}}$ is $\mathcal{M}$-measurable on $A \in \mathcal{M}$, then $f \cdot \mathbf{1}_A$ is $\mathcal{M}$-measurable.*

*Proof.* Letting $f_A = f \cdot \mathbf{1}_A$ and noting that $f_A = f$ over $A$ and $f_A = 0$ over $A^c$, we have for any $E \in \mathcal{B}_{\overline{\mathbb{R}}}$ such that $0 \notin E$,

$$\begin{cases} f_A^{-1}(E) \cap A = f^{-1}(E) \cap A \in \mathcal{M} & \text{(by the } \mathcal{M}\text{-measurability of } f \text{ on } A), \\ f_A^{-1}(E) \cap A^c = \varnothing \in \mathcal{M}. \end{cases}$$

Moreover, it is obvious that $f_A^{-1}\{0\} = A^c \cup \{x \in A : f(x) = 0\}$. Hence,

$$\begin{cases} f_A^{-1}\{0\} \cap A = \{x \in A : f(x) = 0\} = f^{-1}\{0\} \cap A \in \mathcal{M}, \\ f_A^{-1}\{0\} \cap A^c = A^c \in \mathcal{M}. \end{cases}$$

For a general $E \in \mathcal{B}_{\overline{\mathbb{R}}}$, since $f_A^{-1}(E) = f_A^{-1}(E \setminus \{0\}) \cup f_E^{-1}\{0\}$, we have

$$\begin{cases} f_A^{-1}(E) \cap A = \left(f_A^{-1}(E \setminus \{0\}) \cap A\right) \cup \left(f_A^{-1}\{0\} \cap A\right) \in \mathcal{M}, \\ f_A^{-1}(E) \cap A^c = \left(f_A^{-1}(E \setminus \{0\}) \cap A^c\right) \cup \left(f_A^{-1}\{0\} \cap A^c\right) \in \mathcal{M}. \end{cases}$$

Therefore, $f_A$ is $\mathcal{M}$-measurable by Lemma 1.9 with $B = A^c$. □

A real-valued function $\phi : X \to \mathbb{R}$ is said to be simple, or *a simple function*, if it is a linear combination of indicator functions of sets in $\mathcal{M}$, that is, if $\phi(x) = \sum_{i=1}^n a_j \cdot \mathbf{1}_{E_i}(x)$ for some $a_j \in \mathbb{R}$ and $E_i \in \mathcal{M}$ such that $X = \bigcup_{i=1}^n E_i$. The simple functions have some nice properties as shown below.

**Proposition 1.29.** *If $f_1, f_2 : X \to \mathbb{R}$ are simple, then so are $f_1 + f_2$ and $f_1 f_2$. Moreover, $f : X \to \mathbb{R}$ is simple iff $f$ is measurable and the range of $f$ is a finite subset of $\mathbb{R}$.*

*Proof.* $f_1 + f_2$ is simple by the definition of a simple function. $f_1 f_2$ is also simple by the definition and the fact that $\mathbf{1}_E \cdot \mathbf{1}_F = \mathbf{1}_{E \cap F}$ for any $E, F \in X$. Moreover, if $f$ is simple, then it is measurable by Lemma 1.10 and the recursive applications of Proposition 1.25; $f$ has a finite range in $\mathbb{R}$ since $f(x) = \sum_{i=1}^n a_i \mathbf{1}_{E_i}(x)$ is a linear combination of 0's and 1's, uniformly having at most $2^n$ values. Conversely, suppose $f$ is measurable and range$(f) = \{b_1, b_2, \cdots, b_n\}$. Then, for each $i, j \in \{1, 2, \cdots, n\}$, $E_i \doteq f^{-1}\{b_i\} \in \mathcal{M}$ and by the definition of a function, $E_i \cap E_j = \varnothing$. In summary, $f$ can be represented as

$$f = \sum_{i=1}^n b_i \cdot \mathbf{1}_{E_i}, \text{ where } E_i = f^{-1}\{b_i\} \text{ and range}(f) = \{b_1, b_2, \cdots, b_n\}. \tag{1.14}$$

Therefore, $f$ is simple and the proof is completed. □

Here, Proposition 1.29 ensures that every simple function $f$ can be represented by (1.14) called the *standard representation* of $f$, with $\{E_i\}_{i=1}^n$ disjoint. Next, we prove that any extended real-valued measurable function can be approximated by a simple function represented by (1.14) for some $E_i$'s, $b_j$'s and $n$.

**Theorem 1.1.** *If $f : X \to [0, \infty]$ is measurable on a measurable space $(X, \mathcal{M})$, then there exists a sequence $\{\phi_n\}$ of simple functions such that $0 \le \phi_1 \le \phi_2 \le \cdots \le f$ and*

$$\phi_n \to f \text{ as } n \to \infty, \text{ pointwisely on } X, \text{ and uniformly on any subset on which } f \text{ is bounded.} \tag{1.15}$$

*Proof.* Consider the functions $\phi_n : X \to [0, \infty]$ of the form:

$$\phi_1(x) = \begin{cases} 0 \text{ if } f(x) \in [0, 1), \\ 1 \text{ if } f(x) \ge 1, \end{cases} \qquad \phi_2(x) = \begin{cases} 0 \text{ if } f(x) \in [0, \frac{1}{2}), & \frac{1}{2} \text{ if } f(x) \in [\frac{1}{2}, 1), \\ 1 \text{ if } f(x) \in [1, \frac{3}{2}), & \frac{3}{2} \text{ if } f(x) \in [\frac{3}{2}, 2), \\ 2 \text{ if } f(x) \ge 2, \end{cases}$$

$$\phi_3(x) = \begin{cases} 0 \text{ if } f(x) \in [0, \frac{1}{4}), & \frac{1}{4} \text{ if } f(x) \in [\frac{1}{4}, \frac{1}{2}), & \frac{1}{2} \text{ if } x \in [\frac{1}{2}, \frac{3}{4}), & \frac{3}{4} \text{ if } x \in [\frac{3}{4}, 1), \\ 1 \text{ if } f(x) \in [1, \frac{5}{4}), & \frac{5}{4} \text{ if } f(x) \in [\frac{5}{4}, \frac{3}{2}), & \frac{3}{2} \text{ if } x \in [\frac{3}{2}, \frac{7}{4}), & \frac{7}{4} \text{ if } x \in [\frac{7}{4}, 2), \\ 2 \text{ if } f(x) \in [2, \frac{9}{4}), & \frac{9}{4} \text{ if } f(x) \in [\frac{9}{4}, \frac{5}{2}), & \frac{5}{2} \text{ if } f(x) \in [\frac{5}{2}, \frac{11}{4}), & \frac{11}{4} \text{ if } f(x) \in [\frac{11}{4}, 3), \\ 3 \text{ if } f(x) \in [3, \frac{13}{4}), & \frac{13}{4} \text{ if } f(x) \in [\frac{13}{4}, \frac{7}{2}), & \frac{7}{2} \text{ if } f(x) \in [\frac{7}{2}, \frac{15}{4}), & \frac{15}{4} \text{ if } f(x) \in [\frac{15}{4}, 4), \\ 4 \text{ if } f(x) \ge 4. \end{cases}$$

$$\vdots$$

That is, $\phi_{n+1}(x) = \begin{cases} (k-1) \cdot 2^{-n} \text{ if } f(x) \in [(k-1) \cdot 2^{-n}, k \cdot 2^{-n}) \text{ for } k = 1, 2, 3, \cdots, 2^{2n}, \\ \qquad 2^n \quad \text{ if } f(x) \ge 2^n. \end{cases}$ Then, by the above formulas, we can clearly see that $\phi_n \le \phi_{n+1}$ and $0 \le f - \phi_{n+1}$ for all $n$. Moreover, for any $x \in X$ satisfying $f(x) \le 2^n$ and any $n$, we have $f(x) \in [(k-1) \cdot 2^{-n}, k \cdot 2^{-n})$ for some $k \in \{1, 2, 3, \cdots, 2^{2n}\}$, and thereby,

$$0 \le f(x) - \phi_{n+1}(x) = f(x) - (k-1) \cdot 2^{-n} \le 2^{-n}. \tag{1.16}$$

Therefore, $\phi_n(x) \to f(x)$ as $n \to \infty$ for any $x$, pointwise convergence on $X$. Moreover, for any bounded region $E \subseteq \{x \in X : f(x) \leq c\}$ $(c > 0)$, one has (1.16) uniformly in $x \in E$ if $2^n \geq c$ and thus the convergence is uniform on any subset of $X$ on which $f$ is bounded. Finally, the proof is completed as each $\phi_n$ is simple and can be represented as

$$\phi_{n+1} = \sum_{k=1}^{2^{2n}} (k-1)\cdot 2^{-n} \cdot \mathbf{1}_{E_{n,k}} + 2^n \cdot \mathbf{1}_{F_n}, \text{ where } E_{n,k} = f^{-1}\big([(k-1)\cdot 2^{-n}, k\cdot 2^{-n})\big), \ F_n = f^{-1}\big([2^n, \infty]\big).$$

□

**Corollary 1.12.** *If $f : X \to \overline{\mathbb{R}}$ is measurable, then there exists a sequence $\{\phi_n\}$ of simple functions such that $0 \leq |\phi_1| \leq |\phi_2| \leq \cdots \leq |f|$ and (1.15) holds.*

*Proof.* Since $f$ can be decomposed as $f = f^+ - f^-$, where $f^+$ and $f^-$ given by (1.12) are measurable as discussed above, By Theorem 1.1, there are the sequences $\{\phi_n^+\}$ and $\{\phi_n^-\}$ of simple functions such that $0 \leq \phi_1^+ \leq \phi_2^+ \leq \cdots \leq f^+$ and $0 \leq \phi_1^- \leq \phi_2^- \leq \cdots \leq f^-$, both of which implies, when $\phi_n \doteq \phi_n^+ - \phi_n^-$, that $0 \leq |\phi_1| \leq |\phi_2| \leq \cdots \leq |f|$. Here, $\phi_n$'s are also simple by Proposition 1.29 and $|\phi_n| = \phi^+ + \phi^-$; also note that $f = f^+ + f^-$. Finally, the convergence (1.15) can be also established by Theorem 1.1 as $\phi_n^+ \to f^+$ and $\phi_n^- \to f^-$ and hence,

$$\phi_n = \phi_n^+ - \phi_n^- \longrightarrow f^+ - f^- = f \text{ as } n \to \infty,$$

pointwisely on $X$ and uniformly on any subset of $X$ on which $f$ is bounded. □

# Chapter 2

# Measures and Integrations

In this chapter, $(X, \mathcal{M})$ denotes a measurable space and $(X, \mathcal{M}, \mu)$ a measure space (see the definition below). We say that a sequence $\{E_i\}_{i=1}^{\infty}$ in $\mathcal{M}$ is disjoint if $E_i \cap E_j = \varnothing$ for all $i, j \in \mathbb{N}$.

## 2.1 Measures

A measure on $\mathcal{M}$ (or on $(X, \mathcal{M})$ or simply on $X$ if $\mathcal{M}$ is understood) is a function $\mu : \mathcal{M} \to [0, \infty]$ such that

1. $\mu(\varnothing) = 0$,
2. **(countable additivity)** if $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$ is disjoint, then $\mu\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \sum_{i=1}^{\infty} \mu(E_i)$.

The space $(X, \mathcal{M}, \mu)$ is called a measure space. The followings are some types and examples of the measures.

- A measure $\mu$ is said to be *finite* if $\mu(X) < \infty$. In this case, $\mu(E) < \infty$ for every $E \in \mathcal{M}$ since $0 \leq \mu(X) = \mu(E) + \mu(E^c) < \infty$ and $\mu(E), \mu(E^c) \geq 0$. Some examples are as follows.
  - A probability measure $\mathbb{P} : \mathcal{M} \to [0, 1]$, a measure $\mathbb{P}$ with the property $\mathbb{P}(X) = 1$.
  - A dirac measure $\delta_{x_0} : \mathcal{M} \to \{0, 1\}$ defined on a $\sigma$-algebra $\mathcal{M}$ such that $\{x_0\} \in \mathcal{M}$ as
  $$\delta_{x_0}(E) \doteq \mathbf{1}_{E \cap \{x_0\}} \text{ for } E \in \mathcal{M}.$$
  Here, note that a dirac measure is a probability measure.
  - A *finite* Borel measure $\mu : \mathcal{M} \to [0, \infty)$ on $\mathbb{R}$, with $\mathcal{M} = \mathcal{B}_{\mathbb{R}}$ and $\mu(\mathbb{R}) < \infty$. An explicit formula w.r.t. an interval $(a, b]$ can be given by $\mu((a, b]) = F(b) - F(a)$ for the associated distribution function $F(x) \doteq \mu\big((-\infty, x]\big)$. Note that if $\mu(\mathbb{R}) = 1$, then $\mu$ is a probability measure on $\mathbb{R}$ and $F$ is the cumulative distribution function in probability theory (see Section 3.2 for general cases).
- A measure $\mu$ is said to be $\sigma$-finite if there exists $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$ such that
$$X = \bigcup_{i=1}^{\infty} E_i \text{ and } \mu(E_i) < \infty \text{ for all } i \in \mathbb{N}.$$
A finite measure is $\sigma$-finite, but *not* vice versa. Most measures arising in practice are $\sigma$-finite including Lebesgue-Stieltjes measures $\bar{\mu}_F$ in Section 3.2 and the Lebesque measure $\mathfrak{m}$.

The fundamental properties of measures can be summarized in the following theorem.

**Theorem 2.1.** *Let $(X, \mathcal{M}, \mu)$ be a measure space.*

*a.* ***(Monotonicity and Subtractivity)*** *for any $E, F \in \mathcal{M}$, $E \subseteq F \implies \begin{cases} \mu(E) \leq \mu(F) \\ \mu(F \setminus E) = \mu(F) - \mu(E) \end{cases}$*

*Moreover, the followings are true for any sequence* $\{E_i\}_{i=1}^\infty \subseteq \mathcal{M}$.

b. ***(Subadditivity)*** $\mu\Big(\bigcup_{i=1}^\infty E_i\Big) \le \sum_{i=1}^\infty \mu(E_i)$.

c. ***(Continuity from below)*** *If* $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$, *then* $\mu\Big(\bigcup_{i=1}^\infty E_i\Big) = \lim_{i\to\infty} \mu(E_i)$.

d. ***(Continuity from above)*** *If* $E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots$ *and* $\mu(E_1) < \infty$, *then* $\mu\Big(\bigcap_{i=1}^\infty E_i\Big) = \lim_{i\to\infty} \mu(E_i)$.

e. ***(Borel-Cantalli Lemma)*** *If* $\sum_{i=1}^\infty \mu(E_i) < \infty$, *then* $\mu\Big(\limsup_{i\to\infty} E_i\Big) = \mu\Big(\bigcap_{i=1}^\infty \bigcup_{j=i}^\infty E_j\Big) = 0$.

*Proof.* **(Monotonicity and Subtractivity)** If $E, F \in \mathcal{M}$ and $E \subseteq F$, then

$$\begin{aligned}\mu(F) = \mu(E \cup (F \setminus E)) &= \mu(E) + \mu(F \setminus E) \\ &\ge \mu(E).\end{aligned} \tag{2.1}$$

**(Subadditivity and Continuity from below)** Let $F_1 = E_1$ and $F_k = E_k \setminus \big(\bigcup_{i=1}^{k-1} E_i\big)$ for $k \ge 2$. Then, $\{F_k\}_{k=1}^\infty \subseteq \mathcal{M}$ by Proposition 1.1, is disjoint, and satisfies $\bigcup_{i=1}^\infty E_i = \bigcup_{i=1}^\infty F_i$. Therefore, by (a) monotonicity,

$$\mu\Big(\bigcup_{i=1}^\infty E_i\Big) = \mu\Big(\bigcup_{i=1}^\infty F_i\Big) = \sum_{i=1}^\infty \mu(F_i) \le \sum_{i=1}^\infty \mu(E_i).$$

Moreover, let $E_0 \doteq \varnothing$. Then, if $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$, then $F_k = E_k \setminus E_{k-1}$ for all $k$. Hence,

$$\mu\Big(\bigcup_{i=1}^\infty E_i\Big) = \mu\Big(\bigcup_{i=1}^\infty E_i \setminus E_{i-1}\Big) = \sum_{i=1}^\infty \mu(E_i \setminus E_{i-1}) = \lim_{n\to\infty} \sum_{i=1}^n \mu(E_i \setminus E_{i-1}) = \lim_{n\to\infty} \mu(E_n).$$

**(Continuity from above)** Let $F_i \doteq E_1 \setminus E_i$ for $i \in \mathbb{N}$. Then, $F_{i+1} = E_1 \cap E_{i+1}^c \subseteq E_1 \cap E_i^c = F_i$ by $E_{i+1} \supseteq E_i$ for all $i$ and hence, by (c) continuity from below,

$$\mu\Big(E_1 \setminus \Big(\bigcap_{i=1}^\infty E_i\Big)\Big) = \mu\Big(\bigcup_{i=1}^\infty F_i\Big) = \lim_{i\to\infty} \mu(F_i), \tag{2.2}$$

where we have substituted $E_1 \setminus \big(\bigcap_{i=1}^\infty E_i\big) = E_1 \cap \big(\bigcup_{i=1}^\infty E_i^c\big) = \bigcup_{i=1}^\infty E_1 \cap E_i^c = \bigcup_{i=1}^\infty E_1 \setminus E_i = \bigcup_{i=1}^\infty F_i$ for the first equality. Moreover, since $\bigcap_{i=1}^\infty E_i \subseteq E_1$ and $\mu(E_1) = \mu(E_i \cup (E_1 \setminus E_i)) = \mu(E_i) + \mu(F_i)$,

$$\begin{aligned}
\mu(E_1) &= \mu\Big(\bigcap_{i=1}^\infty E_i\Big) + \mu\Big(E_1 \setminus \Big(\bigcap_{i=1}^\infty E_i\Big)\Big) && \text{(by (2.1) with } E = \textstyle\bigcap_{i=1}^\infty E_i \text{ and } F = E_1) \\
&= \mu\Big(\bigcap_{i=1}^\infty E_i\Big) + \lim_{j\to\infty} \mu(F_j) && \text{(by (2.2))} \\
&= \mu\Big(\bigcap_{i=1}^\infty E_i\Big) + \mu(E_1) - \lim_{j\to\infty} \mu(E_i) && \text{(by } \mu(E_1) = \mu(E_i) + \mu(F_j)).
\end{aligned}$$

Finally, since $\mu(E_1) < \infty$, subtracting it from the above equation yields the desired result.

**(Borel-Cantalli Lemma)** Let $S_n = \sum_{i=1}^n \mu(E_i)$ and $S = \lim_{n\to\infty} S_n$. Then, the series $S_n$ converges to $S$ as $n \to \infty$ and thus we have for each $k \in \mathbb{N}$, $\lim_{n\to\infty} \sum_{i=k}^n \mu(E_i) = \lim_{n\to\infty} \big(S_n - S_k\big) = S - S_k$. Hence,

$$\lim_{k\to\infty} \sum_{i=k}^\infty \mu(E_i) = S - S = 0. \tag{2.3}$$

Since $F_k \doteq \bigcup_{i=k}^\infty E_i$ satisfies $F_1 \supseteq F_2 \supseteq F_3 \supseteq \cdots$ and $\mu(F_1) \le \sum_{i=1}^n \mu(E_i) < \infty$ by (b) subadditivity and assumption, we finally obtain by (d) continuity from above, (b) subadditivity, and (2.3) that

$$\mu\Big(\limsup_{i\to\infty} E_i\Big) = \mu\Big(\bigcap_{k=1}^\infty \bigcup_{i=k}^\infty E_i\Big) = \mu\Big(\bigcap_{k=1}^\infty F_k\Big) = \lim_{k\to\infty} \mu(F_k) \le \lim_{k\to\infty} \sum_{i=k}^\infty \mu(E_i) = 0,$$

which completes the proof. $\square$

**Corollary 2.1.** $\mu\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \lim_{n\to\infty} \mu\Big(\bigcup_{i=1}^{n} E_i\Big)$ *for any* $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$.

*Proof.* Applying Theorem 2.1c to $F_n \doteq \bigcup_{i=1}^{n} E_i$, $\mu\Big(\bigcup_{n=1}^{\infty} F_n\Big) = \lim_{n\to\infty} \mu(F_n)$. □

## 2.2 Completeness

A set $E \in \mathcal{M}$ is said to be of measure zero iff $\mu(E) = 0$. From subadditivity in Theorem 2.1, we obtain the following corollary.

**Corollary 2.2.** *Any countable union of measure zero sets is also of measure zero.*

*Proof.* $0 \le \mu\big(\bigcup_{i=1}^{\infty} E_i\big) \le \sum_{i=1}^{\infty} \mu(E_i) = 0$ for any $\{E_i : \mu(E_i) = 0\}_{i=1}^{\infty}$. □

Note that for a measure zero set $E \in \mathcal{M}$ and $F \subset E$, we have $\mu(F) = 0$ by monotonicity provided that $F \in \mathcal{M}$. A measure whose domain includes all subsets of measure zero sets is called *complete*, and the associated measure space is called a *complete* measure space. Completeness can obviates annoying technical points and as expected and shown below can be made by enlarging the domain of the target measure, the $\sigma$-algebra (see the Theorem below and for an extension from outer measures, see Section 3.1).

**Theorem 2.2.** *Let* $\overline{\mathcal{M}} \doteq \{E \cup F : E \in \mathcal{M}$ *and* $F \subseteq N$ *for some measure zero set* $N \in \mathcal{M}\}$ *(here,* $\overline{\mathcal{M}}$ *contains all subsets of measure zero sets in* $\mathcal{M}$ *— consider* $E = \varnothing$*). Then,*

1. $\overline{\mathcal{M}}$ *is a* $\sigma$*-algebra;*
2. $\bar{\mu} : \overline{\mathcal{M}} \to [0, \infty]$ *given by*

$$\bar{\mu}(E \cup F) \doteq \mu(E) \text{ for } E \text{ and } F \text{ in the definition of } \overline{\mathcal{M}} \tag{2.4}$$

   *is the unique extension* $\bar{\mu}$ *of* $\mu$ *to a complete measure on* $\overline{\mathcal{M}}$.

*Proof.* Let $\mathcal{N} \doteq \{N \in \mathcal{M} : \mu(N) = 0\} \subseteq \mathcal{M}$ be the family of measure zero sets. Since $\mathcal{M}$ and $\mathcal{N}$ are closed under countable unions by the definition of a $\sigma$-algebra and Corollary 2.2, respectively, so is $\overline{\mathcal{M}}$ as shown below:

$$\{E_i\}_{i=1}^{\infty} \subset \mathcal{M} \text{ and } \{F_i : F_i \subseteq N_i \text{ for some } N_i \in \mathcal{N}\}_{i=1}^{\infty} \implies \bigcup_{i=1}^{\infty}(E_i \cup F_i) = \big(\textstyle\bigcup_{i=1}^{\infty} E_i\big) \cup \big(\bigcup_{i=1}^{\infty} F_i\big),$$

where $\bigcup_{i=1}^{\infty} E_i \in \mathcal{M}$ and $\bigcup_{i=1}^{\infty} F_i \subseteq \bigcup_{i=1}^{\infty} N_i \in \mathcal{N}$. To show that $\overline{\mathcal{M}}$ is closed under taking complements, suppose $E \cup F \in \overline{\mathcal{M}}$, where $E \in \mathcal{M}$ and $F \subseteq N \in \mathcal{N}$. Without loss of generality, assume that $E \cap N = \varnothing$, hence $E \cap F = \varnothing$ by $F \subseteq N$ (otherwise, replace $F$ and $N$ by $F \setminus E$ and $N \setminus E$, both disjoint from $E$). Then,

$$\begin{aligned} E \cup F &= (E \cup F) \cap \underbrace{(N \cup N^c)}_{X} = \underbrace{(E \cap N)}_{\varnothing} \cup (E \cap N^c) \cup (F \cap N) \cup \underbrace{(F \cap N^c)}_{\varnothing \text{ by } F \subseteq N} = (E \cap N^c) \cup (F \cap N) \\ &= \big[(E \cap N^c) \cup \underbrace{(E \cap F)}_{\varnothing}\big] \cup \big[(F \cap N) \cup \underbrace{(N^c \cap N)}_{\varnothing}\big] = \big[E \cap (N^c \cup F)\big] \cup \big[(N^c \cup F) \cap N\big] \\ &= (E \cup N) \cap (N^c \cup F), \end{aligned}$$

which implies $(E \cup F)^c = (E \cup N)^c \cap (N \setminus F)$. But, $(E \cup N)^c \in \mathcal{M}$ and $N \setminus F \subset N \in \mathcal{N}$, implying $(E \cup F)^c \in \overline{\mathcal{M}}$. Therefore, $\overline{\mathcal{M}}$ is a $\sigma$-algebra.

Next, for $\overline{E} \doteq E \cup F \in \overline{\mathcal{M}}$ for $E \in \mathcal{M}$ and $F \subseteq N \in \mathcal{N}$ as above, we show that $\bar{\mu}(\overline{E}) \doteq \mu(E)$ is a measure on $\overline{\mathcal{M}}$. First, $\bar{\mu}$ is well-defined, since if $\overline{E} = E_1 \cup F_1 = E_2 \cup F_2$ for some $E_1, E_2 \in \mathcal{M}$, $F_1 \subseteq N_1 \in \mathcal{N}$ and $F_2 \subseteq N_2 \in \mathcal{N}$, then

$$\begin{cases} E_1 \subseteq E_1 \cup F_1 \subseteq E_2 \cup N_2 \implies \mu(E_1) \le \mu(E_2) + \mu(N_2) = \mu(E_2) \\ E_2 \subseteq E_2 \cup F_2 \subseteq E_1 \cup N_1 \implies \mu(E_2) \le \mu(E_1) + \mu(N_1) = \mu(E_1), \end{cases}$$

meaning that $\bar{\mu}(\overline{E}) = \mu(E_1) = \mu(E_2)$. Moreover,

1. $\bar{\mu}(\varnothing) = \bar{\mu}(\varnothing \cup \varnothing) = \mu(\varnothing) = 0$;

2. for any $\{\overline{E}_i\}_{i=1}^{\infty} \subset \overline{\mathcal{M}}$ disjoint, there exist $\{E_i\}_{i=1}^{\infty} \subset \mathcal{M}$ and $\{F_i : F_i \subseteq N_i \in \mathcal{N}\}_{i=1}^{\infty} \subset \overline{\mathcal{M}}$ such that $\overline{E}_i = E_i \cup F_i$. Since $E_i \subseteq \overline{E}_i$, $E_i \subseteq \overline{E}_i$, and $\overline{E}_i \cap \overline{E}_j = \varnothing$, $E_i \cap E_j \subseteq \overline{E}_i \cap \overline{E}_j = \varnothing$, that is, $E_i \cap E_j = \varnothing$ and thus $\{E_i\}_{i=1}^{\infty}$ is also disjoint. Since $\bigcup_{i=1}^{\infty} F_i \subseteq \bigcup_{i=1}^{\infty} N_i \in \mathcal{N}$, we have

$$\bar{\mu}\Big(\textstyle\bigcup_{i=1}^{\infty}\overline{E}_i\Big) = \bar{\mu}\Big((\textstyle\bigcup_{i=1}^{\infty}E_i) \cup (\textstyle\bigcup_{i=1}^{\infty}F_i)\Big) = \mu\Big(\textstyle\bigcup_{i=1}^{\infty}E_i\Big) = \displaystyle\sum_{j=1}^{\infty}\mu(E_i) = \sum_{j=1}^{\infty}\bar{\mu}(E_i).$$

To show the completeness of $\bar{\mu}$, suppose $A \subseteq X$ and there is $\overline{E} \in \overline{\mathcal{M}}$ such that $A \subseteq \overline{E}$ and $\bar{\mu}(\overline{E}) = 0$. Then, $\overline{E}$ can be decomposed as $\overline{E} = E \cup F$ for $E \in \mathcal{M}$ and $F \subseteq N \in \mathcal{N}$; since $A \subseteq \overline{E} \subseteq E \cup N \in \mathcal{M}$ with $\mu(E \cup N) = 0$ as shown below:

$$0 \le \mu(E \cup N) \le \mu(E) + \mu(N) = \mu(E) = \bar{\mu}(\overline{E}) = 0,$$

$A$ is a subset of a measure zero set, i.e., $A \in \mathcal{N}$, hence $A = \varnothing \cup A \in \overline{\mathcal{M}}$, meaning that $\bar{\mu}$ is complete.

Finally, suppose $\lambda : \overline{\mathcal{M}} \to [0, \infty]$ is another measure that extends $\mu$, i.e., $\mu(E) = \bar{\mu}(E) = \lambda(E)$ for all $E \in \mathcal{M}$. Let $\overline{E} = E \cup F \in \overline{\mathcal{M}}$ for $\overline{E} \in \mathcal{M}$ and $F \subseteq N \in \mathcal{N}$ as above. Then, since $E \subseteq \overline{E} \subseteq E \cup N \in \mathcal{M}$,

$$\mu(E) = \lambda(E) \le \lambda(\overline{E}) \le \lambda(E \cup N) = \mu(E \cup N) \le \mu(E) + \mu(N) = \mu(E),$$

meaning that $\lambda(\overline{E}) = \mu(E)$ $(= \mu(\overline{E}))$. This shows that $\bar{\mu}$ is the unique extension to a complete measure on $\overline{\mathcal{M}}$, and the proof is completed. □

The measure $\bar{\mu}$ in Theorem 2.2 is called the *completion* of $\mu$, and $\overline{\mathcal{M}}$ is called the *completion* of $\mathcal{M}$ with respect to $\mu$. We also simply call $(X, \overline{\mathcal{M}}, \bar{\mu})$ the *completion* of $(X, \mathcal{M}, \mu)$. In what follows, we denote $(X, \overline{\mathcal{M}}, \bar{\mu})$ the completion of a given measure space $(X, \mathcal{M}, \mu)$ or just a given complete measure space.

We say that a property $P$ about points $x \in X$ is true *$\mu$-almost everywhere (a.e.)* or $P(x)$ is true for *$\mu$-almost every* $x \in X$ iff there exists a measure zero set $N \in \mathcal{M}$ such that

$$N_P \doteq \big\{x \in X : P(x) \text{ is not true}\big\} \subseteq N.$$

Note that the set $N_P$ falsifying the given property $P$ does not necessarily belong to the $\sigma$-algebra $\mathcal{M}$ unless the measure space $(X, \mathcal{M}, \mu)$ is *complete*; since $\mu(N) = 0$, as long as $\mu$ is complete, we have $\mu(N_P) = 0$ by both monotonicity and $N_P \in \mathcal{M}$, for any property $P$. We will also say *a.e.* or for *almost every* $x \in X$, without the prefix "$\mu$-" whenever no confusion exists. In fact, there is no difference between $\mu$-a.e. and $\bar{\mu}$-a.e. as shown below.

**Proposition 2.1.** *A property $P$ is true $\mu$-a.e. iff it is true $\bar{\mu}$-a.e.*

*Proof.* ($\Longrightarrow$) If $P$ is true $\mu$-a.e., then by definition, there exists $N \in \mathcal{M}$ such that $N_P \subseteq N$ and $\mu(N) = 0$. Since $\mathcal{M} \subseteq \overline{\mathcal{M}}$ and $\bar{\mu}$, the completion of $\mu$, is the unique extension of $\mu$ to $\overline{\mathcal{M}}$ by Theorem 2.2, we have $N_P \subseteq N \in \overline{\mathcal{M}}$ and $\bar{\mu}(N) = \mu(N) = 0$. Hence, $P$ holds $\bar{\mu}$-a.e.

($\Longleftarrow$) Suppose $P$ holds $\bar{\mu}$-a.e. Then, $N_P \in \overline{\mathcal{M}}$ and $\bar{\mu}(N_P) = 0$ by completeness; by the definition of $\overline{\mathcal{M}}$ (see Theorem 2.2), there exists $E \in \mathcal{M}$ and $F \subseteq N$ for a measure zero set $N \in \mathcal{M}$ s.t. $N_P = E \cup F$. By the definition of $\bar{\mu}$, we have $0 = \bar{\mu}(N_P) = \mu(E)$, hence it is obvious that $N_P \subseteq E \cup N$ and $\mu(E \cup N) = 0$ by

$$\mu(E \cup N) \le \mu(E) + \mu(N) = 0 + 0 = 0,$$

meaning that $E \cup N$ is a $\mu$-measure zero set, by which $P$ is true $\mu$-a.e. □

In what follows, the fact that a statement is true a.e. means that it is true $\mu$-a.e., $\bar{\mu}$-a.e., and both by Proposition 2.1.

**Proposition 2.2.** *For $\overline{\mathbb{R}}$-valued functions $f$ and $\{f_i\}_{i=1}^{\infty}$, the followings hold iff the measure is complete:*

1. *if $f$ is measurable and $f = \hat{f}$ a.e., then $\hat{f}$ is measurable;*

2. *if $f_i$ is measurable for all $i \in \mathbb{N}$ and $f_i \to f$ a.e., then $f$ is measurable.*

*Proof.* To prove sufficiency of the first part, suppose that $(X, \mathcal{M}, \mu)$ is complete. For the proof of the first part, let $P(x) \doteq$ "$f(x)$ is equal to $\hat{f}(x)$". Then, completeness ensures that $N_P \in \mathcal{M}$ ($\because f = \hat{f}$ a.e.) and thus $N_P^c \in \mathcal{M}$. Next, for any $E \in \mathcal{M}$, since $f$ is measurable and $E \cap N_P^c \in \mathcal{M}$ by Proposition 1.1, we have $f^{-1}(E \cap N_P^c) \in \mathcal{M}$. Hence,

$$\hat{f}^{-1}(E) \cap N_P^c = \hat{f}^{-1}(E \cap N_P^c) = f^{-1}(E \cap N_P^c) \in \mathcal{M},$$

where we have used $f = \hat{f}$ on $N_P^c$ and $E \cap N_P^c \subseteq N_P^c$. On the other hand, $\hat{f}^{-1}(E) \cap N_P$ is a subset of a measure zero set $N_P$ and by completeness belongs to $\mathcal{M}$. Therefore, $\hat{f}$ is measurable by $X = N_P \cup N_P^c$ and Lemma 1.9.

To prove sufficiency of the second part, suppose that $(X, \mathcal{M}, \mu)$ is complete, each $f_i$ is measurable and $f_i \to f$ a.e. Let $\hat{f}$ be defined as $\hat{f}(x) \doteq \limsup_{i\to\infty} f_i(x)$. Then, $\hat{f}$ is measurable by Proposition 1.26 and $\hat{f}(x) = \lim_{i\to\infty} f_i(x)$ for any $x \in X$ for which the limit exists. Therefore, by uniqueness of the limit points and $f_i \to f$ a.e., we have $f = \hat{f}$ a.e. and thus $f$ is measurable by the first statement.

For necessity of the each part, suppose the first or second statement is true and but $\mu$ is not complete, so that there exists a non-measurable set $F \subseteq N$ for some $N \in \mathcal{M}$ such that $\mu(N) = 0$. For the first case, consider $f = \mathbf{1}_N$ and $\hat{f} = \mathbf{1}_F$. Then, $f = \mathbf{1}_N$ is measurable by Lemma 1.10 and $f = \hat{f}$ a.e. So, $\hat{f} = \mathbf{1}_F$ is measurable by the first statement, hence $F \in \mathcal{M}$ again by Lemma 1.10. For the second case, consider $f_i = 0$ and $f = \mathbf{1}_F$, with which we have $f_i \to f$ a.e. (except within the measure zero set $N \supseteq F$), and hence $f = \mathbf{1}_F$ is measurable by the second statement. Therefore, again by Lemma 1.10, $F \in \mathcal{M}$. For both cases, we have proved $F \in \mathcal{M}$, implying that the measure space $(X, \mathcal{M}, \mu)$ is complete. □

If $f$ is $\mathcal{M}$-measurable, then it is $\overline{\mathcal{M}}$-measurable since $\mathcal{M} \subseteq \overline{\mathcal{M}}$ (see Corollary 1.1). The next proposition states that the target function can be made measurable w.r.t. the measure space before completion by redefining it on some measure zero sets.

**Proposition 2.3.** *if $f$ is $\overline{\mathcal{M}}$-measurable, then there exists a $\mathcal{M}$-measurable function $\hat{f}$ such that $f = \hat{f}$ a.e.*

*Proof.* Note that any $\overline{E} \in \overline{\mathcal{M}}$ can be represented as $\overline{E} = E \cup F$ for $E \in \mathcal{M}$ and $F \subseteq N$ for some $N \in \mathcal{M}$ satisfying $\mu(N) = 0$. First, consider a simple function $f = \sum_{i=1}^n a_i \cdot \mathbf{1}_{\overline{E}_i}$ for some sets $\overline{E}_i \in \overline{\mathcal{M}}$ such that $\overline{E}_i = E_i \cup F_i$ for some $E_i \in \mathcal{M}$ and $F_i \subseteq N_i \in \mathcal{M}$ with $\mu(N_i) = 0$; let $\hat{f} = \sum_{i=1}^n a_i \cdot \mathbf{1}_{E_i}$. Then, $f$ and $\hat{f}$ are $\overline{\mathcal{M}}$- and $\mathcal{M}$-measurable by Proposition 1.29, respectively. Moreover, denoting $N \doteq \bigcup_{i=1}^n N_i \in \mathcal{M}$, we have

$$\{x \in X : f(x) \neq \hat{f}(x)\} = \bigcup_{i=1}^n \overline{E}_i \setminus E_i = \bigcup_{i=1}^n F_i \setminus E_i \subseteq \bigcup_{i=1}^n F_i \subseteq N,$$

where $\mu(N) = 0$ by subadditivity: $\mu(N) \le \sum_{i=1}^n \mu(N_i) = 0$. That is, $f = \hat{f}$ a.e.

For the general case, by Corollary 1.12, there exists a sequence of $\overline{\mathcal{M}}$-measurable simple function $\{\phi_n\}$ which converges pointwisely to $f$. As above, let $\psi_n$ be a $\mathcal{M}$-measurable simple function such that $\phi_n = \psi_n$ except on a set $N_n \in \mathcal{M}$ with $\mu(N_n) = 0$. Let $N \doteq \bigcup_{n=1}^\infty N_n$ and $\hat{f} = \limsup_{n\to\infty} \psi_n$ with slight abuse of notations. Then, $\hat{f}$ is $\mathcal{M}$-measurable by Proposition 1.26 and by subadditivity $\mu(N) \le \sum_{n=1}^\infty \mu(N_n) = 0$, we have $\mu(N) = 0$. Moreover, outside the measure zero set $N$,

$$\hat{f} = \limsup_{n\to\infty} \psi_n = \limsup_{n\to\infty} \phi_n = \lim_{n\to\infty} \phi_n = f,$$

meaning that $f = \hat{f}$ a.e., which completes the proof. □

By combining the two propositions above with the fact that the completion always exists by Theorem 2.2, we obtain the following corollary.

**Corollary 2.3.** *If $f_i$ is measurable for all $i \in \mathbb{N}$ and $f_i \to f$ a.e., then there exists a measurable function $\hat{f}$ s.t. $f_i \to \hat{f}$ a.e. (i.e., $f = \hat{f}$ a.e).*

*Proof.* If $f_i$ is $\mathcal{M}$-measurable, then it is $\overline{\mathcal{M}}$-measurable and hence, $f$ is also $\overline{\mathcal{M}}$-measurable by Proposition 2.2. Now, the proof is direct by Proposition 2.3. □

## 2.3 Integration of Non-negative Functions

In this section, we define the integral of a measurable nonnegative extended-real-valued function. To do so, recall that $\mathcal{L}^+$ is the space of all such functions $f : X \to [0, \infty]$ and first consider the integral of a simple function $\phi \in \mathcal{L}^+$. With any of its representation $\phi = \sum_{i=1}^n a_i \cdot \mathbf{1}_{E_i}$ for a disjoint sequence $\{E_i \in \mathcal{M}\}_{i=1}^n$ such that $X = \bigcup_{i=1}^n E_i$ and nonnegative $a_i$'s, we define the integral of $\phi$ as

$$\int \phi \, d\mu \doteq \sum_{i=1}^n a_i \cdot \mu(E_i), \tag{2.5}$$

with the convention that $0 \cdot \infty = 0$ — see (1.8). The convention is essential as the measure inside the summation and thus the integral may equal $\infty$.

For any $A \in \mathcal{M}$, by Proposition 1.29, $\phi \cdot \mathbf{1}_A$ is also simple and thus measurable. Hence, we also define the integral of $\phi$ *over* $A \in \mathcal{M}$ as

$$\int_A \phi \, d\mu \doteq \int \phi \cdot \mathbf{1}_A \, d\mu = \sum_{i=1}^n a_i \cdot \mu(E_i \cap A),$$

where the formula on the right comes from the fact that

$$\begin{aligned} \phi \cdot \mathbf{1}_A &= \sum_{i=1}^n a_i \cdot \mathbf{1}_{E_i} \cdot \mathbf{1}_A \\ &= \sum_{i=1}^n a_i \cdot \mathbf{1}_{E_i \cap A} + a_{n+1} \cdot \mathbf{1}_{A^c} \end{aligned}$$

for $a_{n+1} \doteq 0$ and the *disjoint* sets: $E_i \cap A$ $(j = 1, 2, \cdots, n)$ and $A^c$, all measurable by Proposition 1.1. Note that even if $\mu(A^c) = \infty$, we have $a_{N+1} \cdot \mu(A^c) = 0 \cdot \infty = 0$ by the convention.

When there is no danger of confusion, we shall denote $\int \phi \, d\mu$ by $\int \phi$; whenever necessary, we will also denote it by $\int \phi(x) \, d\mu(x)$ (some authors prefer to write $\int \phi(x) \, \mu(dx)$). In summary: for a simple function $f$,

$$\int_A f \, d\mu = \int_A f = \int_A f(x) \, d\mu(x) = \int f \cdot \mathbf{1}_A \, d\mu, \qquad \int = \int_X .$$

**Proposition 2.4.** *The integral* (2.5) *with respect to a nonnegative simple function $\phi \in \mathcal{L}^+$ is well-defined. That is, if $\phi = \sum_{j=1}^m b_j \cdot \mathbf{1}_{F_j}$ is another representation with a disjoint sequence $\{F_j \in \mathcal{M}\}_{j=1}^m$ such that $X = \bigcup_{j=1}^m F_j$ and nonnegative $b_j$'s, then*

$$\int \phi \, d\mu = \sum_{j=1}^m b_j \cdot \mu(F_j) = \sum_{i=1}^n a_i \cdot \mu(E_i).$$

*Proof.* Note that $1 = \mathbf{1}_X = \sum_{i=1}^n \mathbf{1}_{E_i} = \sum_{j=1}^m \mathbf{1}_{F_j}$ since both $\{E_i\}$ and $\{F_j\}$ are disjoint sequences such that $X = \bigcup_{i=1}^\infty E_i = \bigcup_{j=1}^\infty F_j$. Hence, by $\mathbf{1}_{E_i} \cdot \mathbf{1}_{F_j} = \mathbf{1}_{E_i \cap F_j}$, we have

$$\phi = \sum_{i=1}^n \Big( a_i \mathbf{1}_{E_i} \cdot 1 \Big) = \sum_{i=1}^n \sum_{j=1}^m a_i \cdot \mathbf{1}_{E_i \cap F_j}, \qquad \phi = \sum_{j=1}^m \Big( b_j \mathbf{1}_{F_j} \cdot 1 \Big) = \sum_{i=1}^n \sum_{j=1}^m b_j \cdot \mathbf{1}_{E_i \cap F_j}.$$

where $\{E_i \cap F_j\}_{i,j}$ is disjoint. Now, whenever there is $x \in E_i \cap F_j$, we have $a_i = \phi(x) = b_j$, that is, whenever $E_i \cap F_j \neq \varnothing$, we have $a_i = b_j$ and thus $a_i \cdot \mu(E_i \cap F_j) = b_j \cdot \mu(E_i \cap F_j)$. On the other hand, since $\mu(\varnothing) = 0$, we also have $a_i \cdot \mu(E_i \cap F_j) = b_j \cdot \mu(E_i \cap F_j)$ for the case $E_i \cap F_j = \varnothing$. Hence,

$$\underbrace{\sum_{i=1}^n a_i \cdot \mu(E_i)}_{= \int \phi \, d\mu} = \sum_{i=1}^n a_i \cdot \mu\Bigg( E_i \cap \underbrace{\Big( \bigcup_{j=1}^m F_j \Big)}_{=X} \Bigg) = \sum_{i=1}^n \sum_{j=1}^m \underbrace{a_i \cdot \mu(E_i \cap F_j)}_{= b_j \cdot \mu(E_i \cap F_j)} = \sum_{j=1}^m \Bigg( b_j \cdot \underbrace{\sum_{i=1}^n \mu(E_i \cap F_j)}_{= \mu\left( F_j \cap \left( \bigcup_{i=1}^n E_i \right) \right)} \Bigg) = \sum_{j=1}^m b_j \mu(F_j),$$

which completes the proof. ☐

**Proposition 2.5.** *Let $f, g \in \mathcal{L}^+$ be simple.*

1. $\displaystyle\int c\cdot f = c\cdot\int f$ *for any* $c\in[0,\infty)$,
2. $\displaystyle\int (f+g) = \int f + \int g$,
3. $\displaystyle f\le g \Longrightarrow \int f \le \int g$,
4. *The map* $\displaystyle A\mapsto \int_A f\, d\mu$ *is a measure on* $\mathcal{M}$.

*Proof.* Let $f=\sum_{i=1}^N a_i\cdot \mathbf{1}_{E_i}$ and $g=\sum_{j=1}^M b_j\cdot\mathbf{1}_{F_j}$ be their representations, with their nonnegative constants $a_i$'s and $b_j$'s, and the disjoint sequences $\{E_i\}_{i=1}^n$ and $\{F_j\}_{j=1}^m$ such that $X=\bigcup_{i=1} E_i = \bigcup_{j=1}^\infty F_j$. For the first part, since $c\ge 0$ and $\{E_i\}_{i=1}^n$ is disjoint and covers $X$, it is obvious that

$$\int c\cdot f\, d\mu = \sum_{i=1}^n (c\cdot a_i)\cdot\mu(E_i) = c\cdot\sum_{i=1}^n a_i\cdot\mu(E_i) = c\cdot\int f\, d\mu.$$

by the definition (2.5) of an integral.

For the second and third parts, note that $1=\mathbf{1}_X=\sum_{i=1}^n \mathbf{1}_{E_i} = \sum_{j=1}^m \mathbf{1}_{F_j}$, and $\mathbf{1}_{E_i}\cdot\mathbf{1}_{F_j} = \mathbf{1}_{E_i\cap F_j}$, by which we have

$$\begin{aligned} f &= \sum_{i=1}^n \Big(a_i\cdot\mathbf{1}_{E_i}\cdot 1\Big) = \sum_{i=1}^n\sum_{j=1}^m a_i\cdot\mathbf{1}_{E_i\cap F_j},\\ g &= \sum_{j=1}^m \Big(b_j\cdot\mathbf{1}_{F_j}\cdot 1\Big) = \sum_{i=1}^n\sum_{j=1}^m b_j\cdot\mathbf{1}_{E_i\cap F_j},\\ f+g &= \sum_{i=1}^n\sum_{j=1}^m (a_i+b_j)\cdot\mathbf{1}_{E_i\cap F_j}. \end{aligned} \tag{2.6}$$

Since $\{E_i\cap F_j\}_{i,j}$ is disjoint and $a_i+b_j\ge 0$ for all $i,j$, we have $\displaystyle\int f+g\, d\mu = \sum_{i=1}^n\sum_{j=1}^m (a_i+b_j)\cdot\mu(E_i\cap F_j)$. On the other hand, the finite additivity of $\mu$ implies:

$$\begin{aligned} \int f\, d\mu + \int g\, d\mu &= \sum_{i=1}^n a_i\cdot\mu\Big(E_i\cap\Big(\bigcup_{j=1}^m F_j\Big)\Big) + \sum_{j=1}^m b_j\cdot\mu\Big(F_j\cap\Big(\bigcup_{i=1}^n E_i\Big)\Big)\\ &= \sum_{i=1}^n\sum_{j=1}^m a_i\cdot\mu(E_i\cap F_j) + \sum_{j=1}^m\sum_{i=1}^n b_j\cdot\mu(F_j\cap E_i)\\ &= \sum_{i=1}^n\sum_{j=1}^m (a_i+b_j)\cdot\mu(E_i\cap F_j). \end{aligned}$$

For the proof of the third property, note that $f\le g$ and the representation (2.6) imply that whenever there is $x\in E_i\cap F_j$, we have $a_i=f(x)\le g(x)=b_j$, that is, whenever $E_i\cap F_j\ne\varnothing$, we have $a_i\le b_j$ and thus $a_i\cdot\mu(E_i\cap F_j)\le b_j\cdot\mu(E_i\cap F_j)$. On the other hand, since $\mu(\varnothing)=0$, we also have $a_i\cdot\mu(E_i\cap F_j)=b_j\cdot\mu(E_i\cap F_j)$ for the case $E_i\cap F_j=\varnothing$. Hence,

$$\int f\, d\mu = \sum_{i=1}^n\sum_{j=1}^m a_i\cdot\mu(E_i\cap F_j)\le\sum_{i=1}^n\sum_{j=1}^m b_j\cdot\mu(E_i\cap F_j) = \int g\, d\mu.$$

Finally, if $\{A_k\in\mathcal{M}\}_{k=1}^\infty$ is disjoint and $A=\bigcup_{k=1}^\infty A_k$, then

$$\int_A f\, d\mu = \sum_{i=1}^n a_i\cdot\mu\Big(E_i\cap\underbrace{\Big(\bigcup_{k=1}^\infty A_k\Big)}_{=A}\Big) = \underbrace{\sum_{i=1}^n\sum_{k=1}^\infty a_i\cdot\mu(E_i\cap A_k)}_{\text{by countable additivity}} = \sum_{k=1}^\infty\Big(\underbrace{\sum_{i=1}^n a_i\cdot\mu(E_i\cap A_k)}_{=\int_{A_k} f\, d\mu}\Big) = \sum_{k=1}^\infty\int_{A_k} f\, d\mu.$$

Moreover, $\int_\varnothing f\, d\mu = \sum_{i=1}^n a_i\cdot\mu(E_i\cap\varnothing) = \sum_{i=1}^n a_i\cdot\mu(\varnothing) = 0$ by $\mu(\varnothing)=0$, which completes the proof. □

We will subsequently prove that the properties in Proposition 2.5 are also true and even extended for general nonnegative measurable functions $f, g : X \to [0, \infty]$ (see Proposition 2.7, Theorems 2.4 and 2.5). Here, we extend the integral (2.5) to all functions $f \in \mathcal{L}^+$ by defining

$$\int f \, d\mu \doteq \sup_{0 \le \phi \le f} \int \phi \, d\mu, \tag{2.7}$$

where the supremum is taken with respect to all simple functions $\phi$ in $\mathcal{L}^+$. For a function $f : X \to [0, \infty]$ measurable *on* $E \in \mathcal{M}$, we define the integral of $f$ over $E$ as

$$\int_E f \, d\mu \doteq \int f \cdot \mathbf{1}_E \, d\mu. \tag{2.8}$$

Note that $f$ is definitely measurable on $E$ if it is measurable, but not vice versa (e.g., $f$ should be measurable on both $E$ and some $F \in \mathcal{M}$ such that $E \cup F = X$ — see Lemma 1.9). Proposition 1.28 and the following ensure the integrals above are well-defined.

**Proposition 2.6.** *The integral definitions* (2.5) *and* (2.7) *agree when $f$ is simple.*

*Proof.* Consider the integrals $\int \phi$ and $\int f$ in the sense of (2.5) when $f$ is simple. Then, clearly, we have $0 \le \int f \le \sup_{0 \le \phi \le f} \int \phi$. On the other hand, by the third property of Proposition 2.5 and $0 \le \phi \le f$,

$$0 \le \int \phi \le \int f \quad \text{and thus} \quad 0 \le \sup_{0 \le \phi \le f} \int \phi \le \int f,$$

hence we have $\displaystyle\int f = \sup_{0 \le \phi \le f} \int \phi$. □

**Proposition 2.7.** *For any $f, g \in \mathcal{L}^+$, $\displaystyle\int f \le \int g$ whenever $f \le g$, and $\displaystyle\int c \cdot f = c \cdot \int f$ for all $c \in [0, \infty)$.*

*Proof.* If $f \le g$, it is obvious that $\displaystyle\int f = \sup_{0 \le \phi \le f} \int \phi \le \sup_{0 \le \phi \le g} \int \phi = \int g$. By the linearity of the integral (2.5) for a simple function $\phi$,

$$\int c \cdot f = \sup_{0 \le \phi \le cf} \int \phi = \sup_{0 \le \phi \le f} \int c \cdot \phi = c \cdot \left( \sup_{0 \le \phi \le f} \int \phi \right) = c \cdot \int f \text{ for any } c \in [0, \infty)$$

(for $c = 0$, we can have the same result simply by $\int c \cdot f = \int 0 = 0 = c \cdot \int f$). □

To further investigate the properties and theory of the integral with respect to a measure, we establish one of the fundamental convergence theorem below.

**Theorem 2.3** (Monotone Convergence Theorem)**.** *Let $\{f_n\}$ be a sequence in $\mathcal{L}^+$ such that $f_n \le f_{n+1}$ for all $n$ and $f = \lim_{n\to\infty} f_n$ $(= \sup_{n \in \mathbb{N}} f_n)$, then $\int f = \lim_{n\to\infty} \int f_n$.*

*Proof.* By $f_n \le f_{n+1}$ and Proposition 2.7, we have $\int f_n \le \int f_{n+1}$, implying that $\{ \int f_n \}$ is an increasing sequence in $[0, \infty]$ and thus its limit "$\lim_{n\to\infty} \int f_n \in [0, \infty]$" exists. Similarly, from $f_n \le \sup_{n \in \mathbb{N}} f_n = f$, we also have $\int f_n \le \int f$ for all $n$, where $f$ $(= \sup_{n \in \mathbb{N}} f_n)$ is measurable by Proposition 1.26, and taking the limit $n \to \infty$ yields $\lim_{n\to\infty} \int f_n \le \int f$.

To prove the reverse inequality, fix $\alpha \in (0, 1)$, let $\phi$ be a simple function such that $0 \le \phi \le f$, and $E_n \doteq \{x : \alpha \cdot \phi(x) \le f_n(x)\}$. Then, by Proposition 2.7,

$$\alpha \int_{E_n} \phi \le \int_{E_n} f_n \le \int f_n. \tag{2.9}$$

Since $f_n(x) \le f_{n+1}(x)$ $\forall x \in X$, we have $E_1 \subseteq E_2 \subseteq \cdots \subseteq E_n \subseteq E_{n+1} \subseteq \cdots$; since $f = \lim_{n\to\infty} f_n$, for each $x \in X$, there exists $n \in \mathbb{N}$ such that $\phi(x) \le f_N(x) \le f(x)$ for all $N \ge n$ by the definition of a limit, implying that $\bigcup_{n=1}^{\infty} E_n = X$. Therefore, by the fourth property in Proposition 2.5 and Theorem 2.1c, $\lim_{n\to\infty} \int_{E_n} \phi = \int \phi$, by which and taking the limit $n \to \infty$ of (2.9) we obtain $\alpha \int \phi \le \lim_{n\to\infty} \int f_n$. Since this is true for all $\alpha \in (0, 1)$, it is true for $\alpha = 1$ (consider a sequence $\{\alpha_i \in (0, 1)\}$ converging to 1 and taking the limit $i \to \infty$ on both sides), that is, $\int \phi \le \lim_{n\to\infty} \int f_n$. Taking the supremum, we finally obtain $\int f = \sup_{0 \le \phi \le f} \int \phi \le \lim_{n\to\infty} \int f_n$, which completes the proof. □

Our first application of the monotone convergence theorem (MCT), Theorem 2.3, is the proof of linearity of the integral (2.7) in the next theorem.

**Theorem 2.4** (Linearity of Integrals)**.** *For any $f, g \in \mathcal{L}^+$,*

1. $\displaystyle\int (f+g) = \int f + \int g$,
2. $\displaystyle\int c \cdot f = c \cdot \int f$ *for all* $c \in [0, \infty]$.

*Proof.* By Theorem 1.1, there exists sequences of nonnegative simple functions $\{\phi_n\}$ and $\{\psi_n\}$ that increase to $f$ and $g$, respectively. Then, $\{\phi_n + \psi_n\}$ increases to $f+g$ and thus, by the MCT (Theorem 2.3) and Proposition 2.5, we prove the first part as follows:

$$\int (f+g) = \lim_{n\to\infty} \int (\phi_n + \psi_n) = \lim_{n\to\infty} \int \phi_n + \lim_{n\to\infty} \int \psi_n = \int f + \int g.$$

The second part is obvious by Proposition 2.7 when $c \neq \infty$. For a general case $c \in [0, \infty]$, consider an increasing sequence $\{c_i \in [0, \infty)\}$ converging to $c \in [0, \infty]$. Then, by the MCT and Proposition 2.7 again,

$$\int c \cdot f = \lim_{i\to\infty} \int c_i \cdot f = \Big(\lim_{i\to\infty} c_i\Big) \cdot \int f = c \cdot \int f.$$

□

**Corollary 2.4.** *If $\{f_n\}$ is a finite or infinite sequence in $\mathcal{L}^+$ and $f = \sum_n f_n$, then $\int f = \sum_n \int f_n$.*

*Proof.* By Theorem 2.4, we have $\int \sum_{n=1}^N f_n = \sum_{n=1}^N \int f_n$ for any finite $N \in \mathbb{N}$. Letting $N \to \infty$ and applying the MCT (Theorem 2.3) again, we obtain

$$\sum_{n=1}^{\infty} \int f_n = \lim_{N\to\infty} \sum_{n=1}^{N} \int f_n = \lim_{N\to\infty} \int \sum_{n=1}^{N} f_n = \int \lim_{N\to\infty} \sum_{n=1}^{N} f_n = \int \sum_{n=1}^{\infty} f_n,$$

which completes the proof. □

**Theorem 2.5.** *For $f \in \mathcal{L}^+$, let $\lambda : \mathcal{M} \to [0, \infty]$ be defined as $\lambda(E) \doteq \int_E f \, d\mu$ for $E \in \mathcal{M}$. Then,*

1. *$\lambda$ is a measure on $\mathcal{M}$;*
2. *for any $g \in \mathcal{L}^+$, $\displaystyle\int g \, d\lambda = \int g \cdot f \, d\mu$.*

*Proof.* For any $E \in \mathcal{M}$, $0 \leq \int f \cdot \mathbf{1}_E \, d\mu = \lambda(E)$ by $0 \leq f \cdot \mathbf{1}_E$ and Proposition 2.7. Moreover,

$$\lambda(\varnothing) = \int_{\varnothing} f d\lambda = \int f \cdot \mathbf{1}_{\varnothing} \, d\lambda = \int 0 \, d\mu = 0.$$

If $\{E_i \in \mathcal{M}\}_{i=1}^{\infty}$ is disjoint and $F \doteq \bigcup_{i=1}^{\infty} E_i$, then $F \in \mathcal{M}$ and thus by Proposition 1.25, $f \cdot \mathbf{1}_F$ is $\mathcal{M}$-measurable. By Corollary 2.4 and the definition,

$$\lambda\bigg(\bigcup_{i=1}^{\infty} E_i\bigg) = \int_F f \, d\mu = \int f \cdot \mathbf{1}_F \, d\mu = \int \sum_{i=1}^{\infty} f \cdot \mathbf{1}_{E_i} \, d\mu = \sum_{i=1}^{\infty} \int f \cdot \mathbf{1}_{E_i} \, d\mu = \sum_{i=1}^{\infty} \int_{E_i} f \, d\mu = \sum_{i=1}^{\infty} \lambda(E_i).$$

Therefore, $\lambda$ is a measure on $\mathcal{M}$.

Next, let $g \in \mathcal{L}^+$ be simple and given by $g = \sum_{i=1}^n a_i \cdot \mathbf{1}_{E_i}$ for a disjoint sequence $\{E_i\}_{i=1}^n$ that covers $X$ and nonnegative constants $a_i$'s. Then, by the definitions and Theorem 2.4,

$$\int g \, d\lambda = \sum_{i=1}^{n} a_i \cdot \lambda(E_i) = \sum_{i=1}^{n} a_i \cdot \int_{E_i} f \, d\mu = \sum_{i=1}^{n} a_i \cdot \int f \cdot \mathbf{1}_{E_i} \, d\mu = \int f \cdot \bigg(\sum_{i=1}^{n} a_i \cdot \mathbf{1}_{E_i}\bigg) \, d\mu = \int f \cdot g \, d\mu.$$

If $g \in \mathcal{L}^+$ is not simple, then by Theorem 1.1, there is an increasing sequence $\{g_n\}$ of simple functions $g_n$ converging to $g$ pointwisely. Then, $\{g_n \cdot f\}_{n=1}^{\infty}$ increases pointwisely to $g \cdot f$ and by the MCT,

$$\int g\, d\lambda = \lim_{n\to\infty} \int g_n\, d\lambda = \lim_{n\to\infty} \int g_n \cdot f\, d\mu = \int \lim_{n\to\infty} g_n \cdot f\, d\mu = \int g \cdot f\, d\mu,$$

which completes the proof. □

**Proposition 2.8.** *If $f \in \mathcal{L}^+$, then $\int f = 0$ iff $f = 0$ a.e.*

*Proof.* If $f$ is simple and thus represented as $f = \sum_{i=1}^n a_i \cdot \mu(E_i)$ for nonnegative $a_i$'s, then it is obvious that $\int f = 0$ iff for each $i$, either $a_i = 0$, $\mu(E_i) = 0$, or both, implying $f = 0$ everywhere but within the null set

$$N = \bigcup \big\{E_i : a_i \neq 0\big\}.$$

In general, if $f = 0$ a.e. and $\phi$ is simple with $0 \leq \phi \leq f$, then $\phi = 0$ a.e, hence $\int \phi = 0$. By Theorem 1.1, there exists a sequence of simple functions $\{\phi_n \in \mathcal{L}^+\}$ such that $0 \leq \phi_n \leq f$ and $\phi_n \to f$. Hence, the MCT implies:

$$\int f = \int \lim_{n\to\infty} \phi_n = \lim_{n\to\infty} \int \phi_n = \lim_{n\to\infty} 0 = 0.$$

To prove the converse, consider the sequence $\{E_n\}$ given by $E_n \doteq \{x : f(x) > 1/n\} = f^{-1}((1/n, \infty]) \in \mathcal{M}$ (see Corollary 1.9). Then, its union $E \doteq \bigcup_{n=1}^{\infty} E_n$ also belongs to $\mathcal{M}$ and satisfies

$$E = f^{-1}((0, \infty]) = \big\{x : f(x) > 0\big\}. \tag{2.10}$$

Clearly, $f = 0$ on $E^c$ since $E^c = \big\{x : f(x) = 0\big\}$; "$f = 0$ a.e." implies that $\mu(E) = 0$ by (2.10). Conversely, if it is false that $f = 0$ a.e., then by (2.10), we must have $\mu(E) > 0$, which and the subadditivity in Theorem 2.1 imply:

$$0 < \mu(E) \leq \sum_{n=1}^{\infty} \mu(E_n),$$

meaning that there exists $n \in \mathbb{N}$ such that $\mu(E_n) > 0$. But for any $n \in \mathbb{N}$, we have $f(x) > n^{-1}\ \forall x \in E_n$, implying $f \geq n^{-1} \cdot \mathbf{1}_{E_n}$, and thus the application of Proposition 2.7 yields $\int f \geq n^{-1}\mu(E_n) > 0$. Therefore, the contraposition shows that $\int f = 0$ implies $f = 0$ a.e. □

**Corollary 2.5.** *If $f, g \in \mathcal{L}^+$, then $\int f = \int g$ iff $f = g$ a.e.*

*Proof.* By Proposition 2.8, $\int (f - g) = 0$ iff $f - g = 0$ a.e. iff $f = g$ a.e., and then apply Theorem 2.4. □

**Proposition 2.9.** *For any $f, g \in \mathcal{L}^+$, $\displaystyle\int f \leq \int g$ whenever $f \leq g$ a.e.*

*Proof.* $f \leq g$ a.e. implies that there exists $E \in \mathcal{M}$ such that $f(x) \leq g(x)$ for all $x \in E$ (i.e., $f \cdot \mathbf{1}_E \leq g \cdot \mathbf{1}_E$) and $\mu(E^c) = 0$. Since we trivially have $f \cdot \mathbf{1}_{E^c} = g \cdot \mathbf{1}_{E^c} = 0$ on $E$, it is obvious that $f \cdot \mathbf{1}_{E^c} = g \cdot \mathbf{1}_{E^c} = 0$ a.e. and thus $\int f \cdot \mathbf{1}_{E^c} = \int g \cdot \mathbf{1}_{E^c} = 0$ by Corollary 2.5 and Proposition 2.8. Therefore,

$$\int f = \int f \cdot \mathbf{1}_E + \int f \cdot \mathbf{1}_{E^c} = \int f \cdot \mathbf{1}_E \leq \int g \cdot \mathbf{1}_E = \int g \cdot \mathbf{1}_E + \int g \cdot \mathbf{1}_{E^c} = \int g.$$

by linearity (Theorem 2.4) and Proposition 2.7. □

The generalized MCTs are the following Corollaries, which require $f_n$ to increase (and converge to $f$) a.e., not necessarily pointwisely.

**Corollary 2.6.** *Let $\{f_n\}$ be a sequence in $\mathcal{L}^+$ such that $f_n$ increases to a measurable function $f \in \mathcal{L}^+$ a.e., then $\int f = \lim_{n\to\infty} \int f_n$.*

*Proof.* Suppose $f_n(x)$ increases to $f(x)$ for all $x \in E$, where $\mu(E^c) = 0$. Then obviously, both $f = f \cdot \mathbf{1}_E$ and for all $n$, $f_n = f_n \cdot \mathbf{1}_E$ hold over $E$ and thus hold a.e. by $\mu(E^c) = 0$. Here, $f_n \cdot \mathbf{1}_E$ pointwisely increases to $f \cdot \mathbf{1}_E$. Therefore,

$$\int f \, d\mu = \int f \cdot \mathbf{1}_E \, d\mu = \lim_{n\to\infty} \int f_n \cdot \mathbf{1}_E \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu.$$

by Corollary 2.5 and the MCT. □

**Lemma 2.1.** *If $\{f_n\} \subset \mathcal{L}^+$ converges to $f$ a.e., then there exists $g \in \mathcal{L}^+$ s.t. $f_n \to g$ a.e. (i.e., $f = g$ a.e.). If $f_n \to f$ pointwisely, then $f \in \mathcal{L}^+$.*

*Proof.* By Corollary 2.3, there exists a measurable function $\hat{f} : X \to \overline{\mathbb{R}}$ such that $f_n \to f = \hat{f}$ on some $E \in \mathcal{M}$ such that $\mu(E^c) = 0$. Since the metric space $(\overline{\mathbb{R}}, \bar{d})$ for $\bar{d}$ given by (1.9) is sequentially compact and $[0, \infty] \subset \overline{\mathbb{R}}$ is a compact set, all the limit points of $f_n(x)$, whenever exists, are in $[0, \infty]$, hence $\hat{f}(x) \in [0, \infty]$ for all $x \in E$. Let $g(x) \doteq \max\{\hat{f}(x), 0\}$. Then, $g$ is measurable by Corollary 1.10, $g = \hat{f}$ on $E$ and $g \geq 0$. That is, $g \in \mathcal{L}^+$ and $f_n \to f = \hat{f} = g$ on $E$.

If $f_n \to f$ pointwisely, then $f$ is measurable by Proposition 1.26, and the proof can be done in the same way, but with $E = X$. □

**Corollary 2.7.** *Let $\{f_n\}$ be a sequence in $\mathcal{L}^+$ which increases to a function $f$ a.e., then there exists a measurable function $g \in \mathcal{L}^+$ such that $f = g$ a.e. and $\int g = \lim_{n\to\infty} \int f_n$.*

*Proof.* By Lemma 2.1, there exists a measurable function $g \in \mathcal{L}^+$ such that $f = g$ a.e. (i.e., $f_n$ increases to $g$ a.e.). Now, the proof is obvious by Corollary 2.6. □

All of the MCTs above assume that at least $\{f_n\}$ have to be increasing. If not, the MCTs would not hold even with the convergent sequences, as can be seen with $\mathbf{1}_{(n,n+1)}$ and $n \cdot \mathbf{1}_{(0,1/n)}$, both of which converge to zero, but $\int \mathbf{1}_{(n,n+1)} = \int n \cdot \mathbf{1}_{(0,1/n)} = 1$. The following Fatou's lemma gives an inequality which is valid for any sequence $\{f_n\}$.

**Lemma 2.2** (Fatou's Lemma)**.** *For any sequence $\{f_n\}$ in $\mathcal{L}^+$,*

$$\int \big(\liminf_{n\to\infty} f_n\big) \leq \liminf_{n\to\infty} \int f_n. \tag{2.11}$$

*Proof.* For each $n \in \mathbb{N}$, we have $\inf_{m\geq n} f_m \leq f_k$ for all $k \geq n$, hence $\int \inf_{m\geq n} f_m \leq \int f_k$ for all $k \geq n$ by Proposition 2.7, which again implies $\int \inf_{m\geq n} f_m \leq \inf_{k\geq n} \int f_k$. Now, taking the limit $n \to \infty$, noting the sequence $\{\inf_{m\geq n} f_m\}$ is pointwisely increasing to "$\liminf_{n\to\infty} f_n$" which is measurable by Proposition 1.26, and applying the MCT (Theorem 2.3), we have

$$\int \liminf_{n\to\infty} f_n \, d\mu = \int \lim_{n\to\infty} \Big(\inf_{n\geq m} f_n\Big) d\mu = \lim_{n\to\infty} \int \inf_{n\geq m} f_n \, d\mu \leq \lim_{n\to\infty} \Big(\inf_{n\geq m} \int f_n \, d\mu\Big) = \liminf_{n\to\infty} \int f_n \, d\mu,$$

where we have substituted the definition $\liminf\limits_{n\to\infty} g_n \doteq \lim\limits_{n\to\infty} \Big(\inf\limits_{m\geq n} g_n\Big)$ for any sequence $\{g_n\}_{n=1}^{\infty}$ of functions or real numbers $g_n$. □

Applying Fatou's lemma, we obtain the following for a sequence which is converging but not necessarily monotonically increasing.

**Corollary 2.8.** *If a sequence $\{f_n\} \subset \mathcal{L}^+$ satisfies one of the followings:*

1. *$f_n$ pointwisely converges to $f$, or*

2. *$f_n$ converges to a measurable function $f \in \mathcal{L}^+$ a.e.,*

*then* $\displaystyle\int f \leq \liminf_{n\to\infty} \int f_n$.

*Proof.* Fatou's lemma implies that $f_n$ satisfies (2.11), hence the proof is complete if $\int f = \int \left(\liminf_{n\to\infty} f_n\right)$ and $f \in \mathcal{L}^+$ for each case. For the first case, the former is obvious since $f = \lim_{n\to\infty} f_n = \liminf_{n\to\infty} f_n$; for the latter, apply Lemma 2.1. For the second part, suppose that $f_n \to f$ a.e. and $f \in \mathcal{L}^+$. Then, by $f_n \to f$ a.e., we have $f = \liminf_{n\to\infty} f_n$ a.e., hence $\int f = \int \left(\liminf_{n\to\infty} f_n\right)$ by Corollary 2.5. □

**Corollary 2.9.** *Let $\{f_n\}$ be a sequence in $\mathcal{L}^+$ which converges to $f$ a.e., then there exists a measurable function $g \in \mathcal{L}^+$ such that $f = g$ a.e. and* $\displaystyle\int g \leq \liminf_{n\to\infty} \int f_n$.

*Proof.* Apply Lemma 2.1 and then Corollary 2.8. □

**Proposition 2.10.** *If $f \in \mathcal{L}^+$ and $\int f < \infty$, then $f^{-1}(\{\infty\}) = \{x : f(x) = \infty\}$ is a null-set.*

*Proof.* By $f \in \mathcal{L}^+$, $f$ is measurable and thus $E \doteq f^{-1}(\{\infty\}) \in \mathcal{M}$ ($\because \{\infty\} \in \mathcal{B}_{\overline{\mathbb{R}}}$). If $E$ is not null, then we have a contradiction to $\int f < \infty$ as follows: since $f(x) = \infty$ for $x \in E$ and $f \geq f \cdot \mathbf{1}_E$,

$$\int f \geq \int_E f \ = \infty \cdot \mu(E) = \infty,$$

where $\mu(E) > 0$. Therefore, $E$ must be a null set for $\int f < \infty$. □

## 2.4 Integration of Real-valued Functions

Now, we define the integral of a real-valued measurable functions $f : X \to \mathbb{R}$ by extending the integral introduced in the previous section. In what follows, we say that a real-valued function $f$ is $\mathcal{M}$-measurable iff it is $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurable, or just measurable if $\mathcal{M}$ is understood so that there is no confusion. Note that a real-valued function is identified as

$$\text{a function } f : X \to \overline{\mathbb{R}} \text{ whose image is a subset of } \mathbb{R} \text{ (i.e., } \mathrm{Im}(f) \subseteq \mathbb{R}). \tag{2.12}$$

This clarification defines the inverse map $f^{-1} : \mathcal{P}(\overline{\mathbb{R}}) \to \mathcal{P}(X)$ with the property that

$$f^{-1}(\{\infty\}) = f^{-1}(\{-\infty\}) = \varnothing \tag{2.13}$$

(see also (1.11)). Proposition 1.24 and (2.13) assure that $f$ is $(\mathcal{M}, \mathcal{B}_{\overline{\mathbb{R}}})$-measurable iff it is $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurable, so that both concepts of measurability (i.e., $(\mathcal{M}, \mathcal{B}_{\overline{\mathbb{R}}})$- and $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurability) coincide. In what follows, we call such a function $f$ satisfying (2.12) (and thereby (2.13)) a *real-valued function* $f : X \to \mathbb{R}$, with a slight abuse of notations.

In this note, we construct an integral with respect to a measurable real-valued function. Consider the positive and negative parts of a measurable real-valued function $f$, namely the non-negative real-valued functions $f^+$ and $f^-$ given by (1.12), which are measurable by Corollary 1.11. If at least one of $\int f^+$ and $\int f^-$ is finite, we define

$$\int f \doteq \int f^+ - \int f^-. \tag{2.14}$$

Moreover, for $f$ measurable *on* $E \in \mathcal{M}$, we also define $\displaystyle\int_E f \doteq \int f \cdot \mathbf{1}_E = \int_E f^+ - \int_E f^-$, where the second equality is true by Theorem 2.4 and the definitions (2.8) and (2.14).

**Definition 2.1.** *A function $f : X \to \mathbb{R}$ is integrable iff it is measurable and both $\int f^+$ and $\int f^-$ are finite. Generally, $f$ is integrable on $E \in \mathcal{M}$ iff it is measurable on $E$ and both $\int_E f^+$ and $\int_E f^-$ are finite.*

**Proposition 2.11.** *$f$ is integrable iff $\int |f| < \infty$.*

*Proof.* Suppose that $f$ is integrable. Then, both $\int f^+$ and $\int f^-$ are finite, hence

$$\int |f| = \int f^+ + \int f^- < \infty,$$

by Theorem 2.4 and $|f| = f^+ + f^-$. Conversely, if $0 \leq \int |f| = \int f^+ + \int f^- < \infty$, then trivially $\int f^+ < \infty$ and $\int f^- < \infty$, which completes the proof. □

In what follows, we denote

$$\mathcal{L}^1(X) := \text{the set of integrable real-valued functions on } X;$$

we also denote it by $\mathcal{L}^1(\mu)$, $\mathcal{L}^1(X,\mu)$, or simply $\mathcal{L}^1$ depending on the context.

**Proposition 2.12.** *$\mathcal{L}^1(X)$ is a real vector space and the integral is a linear functional on it. In summary,*

$$\alpha f + \beta g \in \mathcal{L}^1 \text{ and } \int (\alpha f + \beta g) = \alpha \cdot \int f + \beta \cdot \int g \quad \text{for any } \alpha, \beta \in \mathbb{R} \text{ and } f, g \in \mathcal{L}^1.$$

*Proof.* First, by $\int |f| < \infty$ and $\int |g| < \infty$ (both are true by Proposition 2.11) and $|\alpha f + \beta g| \le |\alpha| \cdot |f| + |\beta| \cdot |g|$, we have

$$\int |\alpha f + \beta g| \le |\alpha| \cdot \int |f| \; + \; |\beta| \cdot \int |g| < \infty,$$

which and Proposition 2.11 again imply that $\alpha f + \beta g \in \mathcal{L}^1$. This itself implies that $\mathcal{L}^1$ is a vector space over $\mathbb{R}$ (e.g., for the existence of the identity element $0 \in \mathcal{L}^1$, take $\alpha = \beta = 0$ for any functions $f, g \in \mathcal{L}^1$). Next, by linearity of the integral (Theorem 2.4), it is obvious that for $\alpha > 0$,

$$\int \alpha \cdot f = \int \alpha \cdot f^+ - \int \alpha \cdot f^- = \alpha \cdot \int f^+ - \alpha \cdot \int f^- = \alpha \cdot \left( \int f^+ - \int f^- \right) = \alpha \cdot \int f;$$

for $\alpha = 0$, $\int \alpha \cdot f = \int 0 = 0$ and for $\alpha < 0$,

$$\int \alpha \cdot f = \int (\alpha \cdot f)^+ - \int (\alpha \cdot f)^- = \int (-\alpha f^-) - \int (-\alpha f^+) = \alpha \cdot \left( -\int f^- + \int f^+ \right) = \alpha \cdot \int f.$$

To show additivity, let $h \doteq f + g$ and note that $h^+ - h^- = h = f^+ - f^- + g^+ - g^-$, which implies

$$h^+ + f^- + g^- = h^- + f^+ + g^+.$$

Therefore, by linearity again, we obtain $\int h^+ + \int f^- + \int g^- = \int h^- + \int f^+ + \int g^+$. Finally, regrouping the terms yields the desired result:

$$\int h = \int h^+ - \int h^- = \int f^+ - \int f^- + \int g^+ - \int g^- = \int f + \int g.$$

□

**Proposition 2.13.** *If $f \in \mathcal{L}^1$, then $|\int f| \le \int |f|$.*

*Proof.* $|\int f| = |\int f^+ - \int f^-| \le |\int f^+| + |\int f^-| = \int f^+ + \int f^- = \int |f|$. □

**Proposition 2.14.** *If $f, g \in \mathcal{L}^1$, then $\int_E f = \int_E g$ for all $E \in \mathcal{M}$ iff $\int |f - g| = 0$ iff $f = g$ a.e.*

*Proof.* Let $h \doteq f - g$. Then, $h \in \mathcal{L}^1$ by Proposition 2.12 and thus $\int |h| < \infty$ by Proposition 2.11, meaning that $|h| \in \mathcal{L}^+$. Hence, $\int |f - g| = \int |h| = 0$ iff $h = 0$ a.e. by Proposition 2.8, hence, iff $f = g$ a.e.

Next, suppose $\int |h| = 0$. Then, by Propositions 2.12 and 2.13, we have for each $E \in \mathcal{M}$:

$$\left| \int_E f - \int_E g \right| = \left| \int h \cdot \mathbf{1}_E \right| \le \int |h| \cdot \mathbf{1}_E \le \int |h| = 0,$$

where the last inequality is true by $|h| \cdot \mathbf{1}_E \le |h|$ and Proposition 2.7. Hence, we have $\int_E f = \int_E g$ for each $E \in \mathcal{M}$. Conversely, if $\int_E f = \int_E g$ for all $E \in \mathcal{M}$, then we have $\int_E h = 0$ for all $E \in \mathcal{M}$ by linearity (Proposition 2.12). Take $E = \{x : h(x) \ge 0\}$. Then, since $E \cup E^c = X$ and

$$\begin{aligned} |h| &= |h| \cdot (\mathbf{1}_E + \mathbf{1}_{E^c}) = |h| \cdot \mathbf{1}_E + |h| \cdot \mathbf{1}_{E^c} \\ &= h \cdot \mathbf{1}_E - h \cdot \mathbf{1}_{E^c}, \end{aligned}$$

we finally have

$$\int |h| \; = \int h \cdot \mathbf{1}_E - \int h \cdot \mathbf{1}_{E^c} = \int_E h - \int_{E^c} h \; = 0,$$

which completes the proof. □

Proposition 2.14 shows that for the purpose of integration, it makes no difference if we alter functions on null sets. Indeed, one can integrate functions $f$ that are only defined on a measurable set $E \in \mathcal{M}$ whose complement is null simply by defining $f$ to be zero (or anything else) on $E^c$. In this fashion, we can treat $\overline{\mathbb{R}}$-valued functions that are finite a.e. as real-valued functions for the purpose of integration (see Section 2.5).

In what follows, we show the last of the three basic convergence theorems (the other two being the MCT and Fatou's lemma), the dominated convergence theorem (DCT), and its variants.

**Theorem 2.6** (Dominated Convergence Theorem). *Let $\{f_n\}$ be a sequence in $\mathcal{L}^1$ such that (a) $f_n$ converges to a measurable function $f$ a.e. and (b) there exists a nonnegative $g \in \mathcal{L}^1$ such that $|f_n| \leq g$ a.e. for all $n$. Then, $f \in \mathcal{L}^1$ and $\int f = \lim_{n\to\infty} \int f_n$.*

*Proof.* First, note that:

1. $|f_n| \leq g$ a.e. means that there exists $E_n \in \mathcal{M}$ such that $|f_n(x)| \leq g(x)$ for all $x \in E_n$ and $\mu(E_n^c) = 0$;
2. $f_n \to f$ a.e. implies that there is $F \in \mathcal{M}$ such that $f_n \to f$ poinwisely on $F$ and $\mu(F^c) = 0$.

Let $E \doteq \bigcap_n E_n$ and $H \doteq E \cap F$, both of which belongs to $\mathcal{M}$ by Proposition 1.1. Then, we have $|f_n(x)| \leq g(x)$ for any $n \in \mathbb{N}$ and all $x \in E$, hence for all $x \in H$; taking the limit $n \to \infty$ on both sides yields $|f(x)| \leq g(x)$ for all $x \in H$; by subadditivity and $\mu(E_n^c) = \mu(F^c) = 0$ for all $n$, we have $\mu(H^c) = 0$ as shown below:

$$\mu(H^c) = \mu(E^c \cup F^c) \leq \mu(E^c) + \mu(F^c) \leq \textstyle\sum_n \mu(E_n^c) + \mu(F^c) = 0.$$

Therefore, $|f| \leq g$ a.e. Since $|f| = f^+ + f^-$ is measurable by Corollary 1.11,

$$\int |f| \leq \int g < \infty$$

by Propositions 2.9 and 2.11 with $0 \leq g \in \mathcal{L}^1$. Therefore, $f \in \mathcal{L}^1$ again by Proposition 2.11.

Next, since $|f_n| \leq g$ on $H$ for all $n$, we have $-f_n \leq g$ and $f_n \leq g$ both on $H$; that is,

$$g + f_n \geq 0 \text{ and } g - f_n \geq 0 \text{ both on } H. \tag{2.15}$$

Similarly, $g+f \geq 0$ and $g-f \geq 0$ both on $H$ from $|f| \leq g$ on $H$. It is trivial that $g \cdot \mathbf{1}_{H^c} = f_n \cdot \mathbf{1}_{H^c} = f \cdot \mathbf{1}_{H^c} = 0$ on $H$, meaning that it is true a.e. and thus

$$\int g \cdot \mathbf{1}_{H^c} = \int f_n \cdot \mathbf{1}_{H^c} = \int f \cdot \mathbf{1}_{H^c} = 0$$

by Proposition 2.14. Moreover, $f_n \to f$ on $F$ implies $f_n \to f$ on $H$ and hence,

$$\liminf_{n\to\infty} f_n = \limsup_{n\to\infty} f_n = \lim_{n\to\infty} f_n = f \text{ on } H,$$

where the limit is a.e.-defined (i.e., defined on $H$). Therefore, by linearity (Theorem 2.4 and Proposition 2.12), the application of Fatou's lemma (Lemma 2.2) to (2.15), and the fact that both integrals (2.7) and (2.14) are equivalent for a non-negative real-valued function $f$, we finally have

$$\int g + \int f = \int (g+f) \cdot \mathbf{1}_H \leq \liminf_{n\to\infty} \int (g+f_n) \cdot \mathbf{1}_H = \int g \cdot \mathbf{1}_H + \liminf_{n\to\infty} \int f_n \cdot \mathbf{1}_H = \int g + \liminf_{n\to\infty} \int f_n$$

$$\int g - \int f = \int (g-f) \cdot \mathbf{1}_H \leq \liminf_{n\to\infty} \int (g-f_n) \cdot \mathbf{1}_H = \int g \cdot \mathbf{1}_H - \limsup_{n\to\infty} \int f_n \cdot \mathbf{1}_H = \int g - \limsup_{n\to\infty} \int f_n.$$

Therefore, $\limsup_{n\to\infty} \int f_n \leq \int f \leq \liminf_{n\to\infty} \int f_n$ and the result follows. □

The following is a concatenation of the variants of the DCT (Theorem 2.6) and itself. The issue here is the measurability of the function $f$ to which the sequence $\{f_n\} \subset \mathcal{L}^1$ converges.

**Corollary 2.10.** *Let $\{f_n\}$ be a sequence in $\mathcal{L}^1$ such that* *(a) at least one of the followings is true:*

- $f_n$ *converges to a measurable function* $f$ *a.e.;*
- $f_n$ *converges to* $f$ *pointwisely;*
- $f_n$ *converges to* $f$ *a.e. and the measure space is complete,*

*(b) there exists a nonnegative* $g \in \mathcal{L}^1$ *such that* $|f_n| \le g$ *a.e. for all* $n$.

*Then,* $f \in \mathcal{L}^1$ *and* $\int f = \lim_{n\to\infty} \int f_n$.

*Proof.* If $f_n$ converges to a measurable function $f$ a.e., then the statement becomes exactly same as the DCT (Theorem 2.6). For the other two cases, $f$ is measurable by the respective Propositions 1.25 and 2.2, thereafter the proof is obvious by Theorem 2.6. □

Another corollary is the bounded convergence theorem, which is useful when the measure space is finite.

**Corollary 2.11** (Bounded Convergence Theorem). *Suppose the measure space is finite (i.e.,* $\mu(X) < \infty$*) and* $\{f_n\}$ *is a sequence of uniformly a.e.-bounded real-valued measurable functions which satisfies at least one of the three convergence conditions in Corollary 2.10, then* $f \in \mathcal{L}^1$ *and* $\int f = \lim_{n\to\infty} \int f_n$.

*Proof.* Since the sequence $\{f_n\}$ is uniformly a.e.-bounded, there is a real number $M > 0$ and $E \in \mathcal{M}$ such that $|f_n(x)| \le M$ for all $x \in E$ and all $n$, and $\mu(E^c) = 0$. Let $g(x) \doteq M$ for all $x \in X$. Then, the sequence is dominated by $g$ a.e. (i.e., $|f_n| \le g$ on $E$). Moreover, $g \in \mathcal{L}^1$ since it is a constant function on a set of finite measure (clearly, $\int g\, d\mu = \int M\, d\mu = M \cdot \mu(X) < \infty$). Therefore, the result follows from Corollary 2.10. □

The DCT and the bounded convergence theorem can be extended to the $L^p$-space (see Section 3.5).

## 2.5 Integration of General Extended Read-valued Functions

In this section, we extend the integral defined in the previous sections to that with respect to a class of a.e.-defined extended real-valued functions, which is not necessarily measurable on a null set. The idea behind this is that by Proposition 2.14, any two functions in $\mathcal{L}^1$ that are equal a.e. have the same integral. To be precise, we define the equivalence class $[g]_E$ of a function $g : X \to \overline{\mathbb{R}}$ on $E \subseteq X$ as

$$[g]_E \doteq \big\{f :\ f \text{ is a.e.-defined and } f = g \text{ a.e., on } E\big\} \quad \text{and} \quad [g] \doteq [g]_X.$$

**Definition 2.2.** *An a.e.-defined function* $f : X \to \overline{\mathbb{R}}$ *is integrable iff there exists* $g \in \mathcal{L}^1$ *such that* $f \in [g]$*. Generally,* $f$ *is integrable on* $E \in \mathcal{M}$ *iff there exists* $g : X \to \overline{\mathbb{R}}$ *such that* $g$ *is integrable on* $E$ *and* $f \in [g]_E$.

For an integrable function $f : X \to \overline{\mathbb{R}}$ defined a.e., we define its integral as

$$\int f\, d\mu \doteq \int g\, d\mu \quad \text{for } g \in \mathcal{L}^1 \text{ such that } f \in [g]. \tag{2.16}$$

Furthermore, for $f$ integrable on $E \in \mathcal{M}$ and a set $\overline{E} \doteq E \cup F$ such that $F \subseteq N$ for some $N \in \mathcal{M}$ with $\mu(N) = 0$, we define the integral of $f$ on $\overline{E}$ as

$$\int_{\overline{E}} f\, d\mu \doteq \int_E g\, d\mu = \int g \cdot \mathbf{1}_E\, d\mu$$

for $g : X \to \mathbb{R}$ such that $g$ is integrable on $E$ and $f \in [g]_E$. By Proposition 2.10, a necessary condition for $f$ to be integrable is that $f^{-1}(\{\pm\infty\}) = \{x : f(x) = \pm\infty\}$ is a null set; otherwise, any measurable function $g$ that is equal to $f$ a.e. has $\mu(E) > 0$ for $E \doteq \{x : g(x) = \pm\infty\}$, resulting in the contradiction to $g \in \mathcal{L}^1$:

$$\int |g|\, d\mu \ge \int_E |g|\, d\mu = \infty \cdot \mu(E) = \infty,$$

hence $g \notin \mathcal{L}^1$ by Proposition 2.11. In what follows, we denote

$$\mathcal{L}^1_{\mathrm{e}} = \{f : f \text{ is a.e.-defined and integrable}\},$$

where the subscript 'e' denotes 'extended'.

The properties shown in the previous section with respect to the integral of a real-valued function can be easily extended to this generalized integral. Moreover, considering $\rho$ defined for $f, g \in \mathcal{L}^1_{\mathrm{e}}$ as

$$\rho_{\mathrm{e}}(f,g) \doteq \int |\hat{f} - \hat{g}| \quad \text{for } \hat{f},\, \hat{g} \in \mathcal{L}^1 \text{ such that } f \in [\hat{f}] \text{ and } g \in [\hat{g}] \tag{2.17}$$

the space $(\mathcal{L}^1_{\mathrm{e}}, \rho_{\mathrm{e}})$ constitutues a pseudometric space, where $\rho_{\mathrm{e}}(f,g) = 0$ iff $f = g$ a.e. If one considers the following space $[\mathcal{L}^1]$ with its metric $\kappa$:

$$[\mathcal{L}^1] \doteq \{[f] :\ f \in \mathcal{L}^1\} \text{ and } \kappa([f],[g]) = \int |f - g| \quad \text{for } f, g \in \mathcal{L}^1,$$

then $([\mathcal{L}^1], \kappa)$ is a metric space ($\because$ $[f] = [g]$ iff $f = g$ a.e. iff $\kappa([f],[g]) = \int |f - g| = 0$ by Proposition 2.14). In what follows, we focus on extending the DCT in a simplified manner.

**Theorem 2.7.** *Let $\{f_n\}$ be a sequence in $\mathcal{L}^1_{\mathrm{e}}$ such that (a) $f_n \to f$ a.e. and (b) there exists a nonnegative $g \in \mathcal{L}^1_{\mathrm{e}}$ such that $|f_n| \le g$ a.e. for all $n$. Then, $f \in \mathcal{L}^1_{\mathrm{e}}$ and $\int f = \lim_{n\to\infty} \int f_n$.*

*Proof.* By the definition of $\mathcal{L}^1_{\mathrm{e}}$, there exists $\{\hat{f}_n\} \subset \mathcal{L}^1$ and $\hat{g} \in \mathcal{L}^1$ such that $f_n = \hat{f}_n$ a.e. and $g = \hat{g}$ a.e. Therefore, $f_n \to f$ a.e. and $|f_n| \le g$ a.e. imply that $\hat{f}_n \to f$ a.e. and $|\hat{f}_n| \le \hat{g}$ a.e., respectively. Moreover, by Proposition 2.3 and measurability of $\hat{f}_n$ for each $n$, there exists a measurable function $\hat{f}$ such that $f = \hat{f}$ a.e. and $f_n \to \hat{f}$ a.e. Therefore, by the DCT (Theorem 2.6), $\hat{f} \in \mathcal{L}^1$ and $\int \hat{f} = \lim_{n\to\infty} \int \hat{f}_n$; the result follows by the definition of $\mathcal{L}^1_{\mathrm{e}}$ and the integral (2.16). □

**Theorem 2.8.** *Let $\{f_n\}$ be a sequence in $\mathcal{L}^1_{\mathrm{e}}$ such that $\sum_{j=1}^{\infty} \int |f_j| < \infty$. Then, $\sum_{j=1}^{n} f_j$ converges a.e. to a function in $\mathcal{L}^1_{\mathrm{e}}$ and $\int \sum_{j=1}^{\infty} f_j = \sum_{j=1}^{\infty} \int f_j$.*

*Proof.* For each $n$, there exists $\hat{f}_n \in \mathcal{L}^1$ such that $f_n = \hat{f}_n$ a.e. and thus $|f_n| = |\hat{f}_n|$ a.e. and by the definition (2.16) of an integral and Proposition 2.11, $\int |f_n| = \int |\hat{f}_n| < \infty$. Hence, by Corollary 2.4,

$$\int \sum_{j=1}^{\infty} |\hat{f}_j| \; = \sum_{j=1}^{\infty} \int |\hat{f}_j| \; = \sum_{j=1}^{\infty} \int |f_j| < \infty,$$

meaning $g \doteq \sum_{j=1}^{\infty} |\hat{f}_j| \in \mathcal{L}^1$ (note that $\sum_{j=1}^{\infty} |\hat{f}_j| = \sup_{n\in\mathbb{N}} \sum_{j=1}^{n} |\hat{f}_j|$ is measurable by Proposition 1.26). In particular, by Proposition 2.10, $\sum_{j=1}^{\infty} |\hat{f}_j(x)|$ is finite for a.e. $x$ and for each such $x$ the series $\sum_{j=1}^{n} \hat{f}_j(x)$ converges. Since $\sum_{j=1}^{n} f_j \overset{\text{a.e.}}{=} \sum_{j=1}^{n} \hat{f}_j$ for all $n$, the a.e.-convergence of $\sum_{j=1}^{n} \hat{f}_j(x)$ means that of $\sum_{j=1}^{n} f_j(x)$ and in addition, we have

$$\Big|\sum_{j=1}^{n} f_j\Big| = \Big|\sum_{j=1}^{n} \hat{f}_j\Big| \le \sum_{j=1}^{\infty} |\hat{f}_j| = g \quad \text{a.e.}$$

Therefore, Theorem 2.7 concludes that $\sum_{j=1}^{\infty} f_j \in \mathcal{L}^1_{\mathrm{e}}$ and $\int \sum_{j=1}^{\infty} f_j = \sum_{j=1}^{\infty} \int f_j$. □

For the further discussions, we also define the measurability of an a.e.-defined function as follows.

**Definition 2.3.** *An a.e.-defined function $f : X \to [0,\infty]$ is measurable, denoted by $f \in \mathcal{L}^+_{\mathrm{e}}$, iff there exists $g \in \mathcal{L}^+$ such that $f \in [g]$.*

# Chapter 3

# More on Measures and Integrations

## 3.1 Outer Measures

Outer measure are useful tools for constructing (complete) measures; its definition comes with the monotonicity and subadditivity properties in Theorem 2.1 in Section 2.1 and conceptually means an approximation of a measure from outside.

**Definition 3.1.** *An outer measure $\mu^*$ on a non-empty set $X$ is a function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ such that*

1. $\mu(\varnothing) = 0$,
2. ***(Monotonicity)*** $E \subseteq F \Longrightarrow \mu^*(E) \leq \mu^*(F)$,
3. ***(Subadditivity)*** $\mu^*\Big(\bigcup_{i=1}^{\infty} E_i\Big) \leq \sum_{j=1}^{\infty} \mu^*(E_i)$ , *where* $\{E_i\}_{i=1}^{\infty}$ *can be disjoint or not.*

The domain $\mathcal{P}(X)$ of an outer measure is the largest $\sigma$-algebra on $X$. A most common way of constructing outer measures is to start with a family $\mathcal{E}$ of "elementary sets" satisfying certain measure-theoretic properties shown below.

**Proposition 3.1.** *Let $\mathcal{E} \subseteq \mathcal{P}(X)$ and $\rho : \mathcal{E} \to [0, \infty]$ be such that $\varnothing, X \in \mathcal{E}$ and $\rho(\varnothing) = 0$. Define*

$$\mu^*(A) = \inf\left\{ \sum_{j=1}^{\infty} \rho(E_j) : E_j \in \mathcal{E} \text{ and } A \subseteq \textstyle\bigcup_{j=1}^{\infty} E_j \right\}$$

*for $A \in \mathcal{P}(X)$. Then, $\mu^*$ is an outer measure on $X$.*

*Proof.* For any $A \in \mathcal{P}(X)$, there exists $\{E_j\}_{j=1}^{\infty} \subseteq \mathcal{E}$ such that $A \subseteq \bigcup_{j=1}^{\infty} E_j$ (take $E_j = X$ for all $j$) so the definition of $\mu^*$ makes sense. If we take $E_j = \varnothing$ for all $j$ and $A = \varnothing \in \mathcal{E}$, we obtain $\mu^*(\varnothing) = 0$. Next, it is obvious that if $A \subseteq B$, then $B \subseteq \bigcup_{j=1}^{\infty} E_j$ implies $A \subseteq \bigcup_{j=1}^{\infty} E_j$, hence

$$\Big\{ \textstyle\sum_{j=1}^{\infty} \rho(E_j) : E_j \in \mathcal{E} \text{ and } B \subseteq \bigcup_{j=1}^{\infty} E_j \Big\} \subseteq \Big\{ \textstyle\sum_{j=1}^{\infty} \rho(E_j) : E_j \in \mathcal{E} \text{ and } A \subseteq \bigcup_{j=1}^{\infty} E_j \Big\},$$

which and the property of the infimum ($S \subseteq T \subseteq \mathbb{R}$ implies $\inf T \leq \inf S$) result in $\mu^*(A) \leq \mu^*(B)$. To prove the countable subadditivity, suppose $\{A_j\}_{j=1}^{\infty} \subseteq \mathcal{P}(X)$ and $\varepsilon > 0$. Then, by the property of the infimum:

$$\begin{aligned} &\text{for every } \delta > 0, \text{ there exists } x \in S \text{ such that } x \leq \inf S + \delta, \\ &\text{(in our case, } S \subseteq [0, \infty] \text{ and the equality explains the case } \infty = \infty), \end{aligned} \tag{3.1}$$

for each $j$, there exists $\{E_j^k\}_{k=1}^{\infty} \subseteq \mathcal{E}$ such that $A_j \subseteq \bigcup_{k=1}^{\infty} E_j^k$ and $\sum_{k=1}^{\infty} \rho(E_j^k) \leq \mu^*(A_j) + \varepsilon \cdot 2^{-j}$; summing up for all $j$ results in

$$\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \rho(E_j^k) \leq \sum_{j=1}^{\infty} \mu^*(A_j) + \varepsilon. \tag{3.2}$$

Since $\bigcup_{j=1}^{\infty} A_j \subseteq \bigcup_{j,k=1}^{\infty} E_j^k$, taking the infimum on (3.2) for all such sets $\{E_j^k \in \mathcal{E}\}_{j,k=1}^{\infty}$ and then limiting $\varepsilon \to 0$ (as $\varepsilon$ is arbitrary) results in $\mu^*(\bigcup_{j=1}^{\infty} A_j) \leq \sum_{j=1}^{\infty} \mu^*(A_j)$. □

If $\mu^*$ is an outer measure on $X$, a set $A \subseteq X$ is called *$\mu^*$-measurable* iff

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ for all } E \subseteq X.$$

Of course, by subadditivity, the inequality $\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c)$ holds for any $A$ and $E$, so to prove that $A$ is $\mu^*$-measurable, it suffices to prove the reverse inequality, which is trivial if $\mu^*(E) = \infty$. In summary, $A$ is $\mu^*$-measurable iff:

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ for all } E \subseteq X \text{ such that } \mu^*(E) < \infty.$$

The motivation for the notion of $\mu^*$-measurability is that for $E \subseteq A$, $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$ means that the outer measure of $A$, $\mu^*(A)$, is equal to the inner measure $\mu^*(E) - \mu^*(E \setminus A)$. The following theorem formalizes this idea and extends it to general $\mu^*$-measurable sets.

**Theorem 3.1** (Carathéodory's Theorem)**.** *If $\mu^*$ is an outer measure on $X$, then the collection $\mathcal{M}$ of all $\mu^*$-measurable sets is a $\sigma$-algebra, and the restriction of $\mu^*$ to $\mathcal{M}$ is a complete measure.*

*Proof.* First, since the definition of $\mu^*$-measurability is symmetric in $A$ and $A^c$, $\mathcal{M}$ is closed under taking complements. Next, for $A, B \in \mathcal{M}$ and $E \subseteq X$,

$$\begin{aligned}\mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \\ &= \underbrace{\mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c)}_{=\mu^*(E\cap A)} + \underbrace{\mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c)}_{=\mu^*(E\cap A^c)}\end{aligned}$$

by the definition of $\mu^*$-measurability. On the other hand, $A \cup B = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B)$, so by subadditivity of $\mu^*$,

$$\mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) \geq \mu^*(E \cap (A \cup B))$$

and hence, we obtain

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c),$$

meaning that $A \cup B$ is $\mu^*$-measurable and thus belongs to $\mathcal{M}$. Therefore, $\mathcal{M}$ is an algebra. Moreover, if $A, B \in \mathcal{M}$ and $A \cap B = \varnothing$, then since $A$ and $B$ are disjoint,

$$\begin{aligned}\mu^*(A \cup B) &= \mu^*((A \cup B) \cap A) + \mu^*((A \cup B) \cap A^c) \\ &= \mu^*(A \cup (B \cap A)) + \mu^*(\varnothing \cup (B \setminus A)) = \mu^*(A) + \mu^*(B),\end{aligned}$$

so $\mu^*$ is finitely additive on $\mathcal{M}$.

To show that $\mathcal{M}$ is a $\sigma$-algebra, let $\{A_j\}_{j=1}^\infty$ be a sequence of disjoint sets in $\mathcal{M}$, $B_n \doteq \bigcup_{j=1}^n A_j$ and $B = \bigcup_{j=1}^\infty A_j$. Then for any $E \subseteq X$, since $\{A_j\}_{j=1}^\infty$ is disjoint, we have $B_n \cap A_n = A_n$ and $B_n \setminus A_n = B_{n-1}$ for all $n$, and thereby,

$$\begin{aligned}\mu^*(E \cap B_n) = \mu^*(E \cap B_n \cap A_n) + \mu^*(E \cap B_n \cap A_n^c) &= \mu^*(E \cap A_n) + \mu^*(E \cap B_{n-1}) \\ &= \mu^*(E \cap A_n) + \mu^*(E \cap A_{n-1}) + \mu^*(E \cap B_{n-2}) \\ &\ \vdots \\ &= \textstyle\sum_{j=1}^n \mu^*(E \cap A_j) \qquad \forall n \in \mathbb{N}.\end{aligned}$$

Moreover, it is obvious by subadditivity that $\mu^*(E \setminus B_n) \geq \mu^*(E \setminus B)$, hence we obtain

$$\mu^*(E) = \mu^*(E \cap B_n) + \mu^*(E \cap B_n^c) \geq \sum_{j=1}^n \mu^*(E \cap A_j) + \mu^*(E \cap B^c)$$

and letting $n \to \infty$ and using subadditivity results in

$$\begin{aligned}\mu^*(E) &\geq \sum_{j=1}^\infty \mu^*(E \cap A_j) + \mu^*(E \cap B^c) \geq \mu^*\left(\textstyle\bigcup_{j=1}^\infty (E \cap A_j)\right) + \mu^*(E \cap B^c) \\ &= \mu^*(E \cap B) + \mu^*(E \cap B^c) \geq \mu^*(E),\end{aligned}$$

where all the inequalities thus become equalities. Therefore, we have $B \in \mathcal{M}$ and thus $\mathcal{M}$ is a $\sigma$-algebra. Moreover, taking $E = B$ results in $\mu^*(B) = \sum_{j=1}^{\infty} \mu^*(A_j)$, implying that $\mu^*|_{\mathcal{M}}$ is countably additive and thus a measure on $(X, \mathcal{M})$.

Finally, to show the completeness of $\mu^*$, it suffices to prove that $\mu^*(A) = 0 \Longrightarrow A \in \mathcal{M}$. Obviously, if $\mu^*(A) = 0$, then for any $E \subseteq X$, we have by the properties of an outer measure,

$$\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c) \leq \mu^*(A) + \mu^*(E \cap A^c) = \mu^*(E \cap A^c) \leq \mu^*(E), \tag{3.3}$$

so that $A \in \mathcal{M}$. Therefore, the measure $\mu^*|_{\mathcal{M}}$ is complete. □

The Carathéodory's Theorem can be used to extend the domains of a premeasure $\mu_0 : \mathcal{A} \to [0, \infty]$ and a measure $\mu : \mathcal{M} \to [0, \infty]$ to construct a measure from $\mu_0$ and a complete measure from $\mu$ on their extended domains $\mathcal{M}$ and $\overline{\mathcal{M}}$, respectively, where $\mathcal{A}$ is an algebra, $\mathcal{M}$ is a $\sigma$-algebra (in the former, $\mathcal{M} \doteq \sigma(\mathcal{A})$), and $\overline{\mathcal{M}}$ is the completion of $\mathcal{M}$. We first consider the former case below.

**Definition 3.2.** *For an algebra $\mathcal{A} \subseteq \mathcal{P}(X)$, a map $\mu_0 : \mathcal{A} \to [0, \infty]$ is said to be a* premeasure *iff:*

- $\mu_0(\varnothing) = 0$;
- *if $\{A_j\}_{j=1}^{\infty} \subseteq \mathcal{A}$ is disjoint and $\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}$, then $\mu_0(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mu_0(A_j)$.*

Note that a premeasure is the same as a measure except that its domain is an algebra, not necessarily a $\sigma$-algebra. Hence, the countable additivity holds only when $\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}$ in its definition above. In particular, a premeasure is finitely additive since one can take $A_j = \varnothing$ for $j$ large. Theorem 2.1 also holds even if $\mu$ is a premeasure, provided that its domain contains the respective underlying countable unions and intersections in the properties.

**Lemma 3.1** (Monotonicity)**.** *$E \subseteq F \Longrightarrow \mu_0(E) \leq \mu_0(F)$ for any $E, F \in \mathcal{A}$.*

*Proof.* $\mu_0(F) = \mu_0(E \cup (F \setminus E)) = \mu_0(E) + \mu_0(F \setminus E) \geq \mu_0(E)$. □

A finite and a $\sigma$-finite premeasure are defined in a similar manner to a finite and a $\sigma$-finite measure, respectively — a premeasure $\mu_0$ is said to be *finite* iff $\mu_0(X) < \infty$ and $\sigma$-finite iff there exists $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{A}$ such that $X = \bigcup_{i=1}^{\infty} E_i$ and $\mu_0(E_i) < \infty$ for all $i \in \mathbb{N}$.

If $\mu_0$ is a premeasure on $\mathcal{A} \subseteq \mathcal{P}(X)$, then it induces an outer measure $\mu^*$ in accordance with Proposition 3.1 ($\varnothing, X \in \mathcal{A}$ by Proposition 1.1) given by

$$\mu^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu_0(A_j) : A_j \in \mathcal{A}, E \subseteq \bigcup_{j=1}^{\infty} A_j \right\} \text{ for any } E \subseteq X. \tag{3.4}$$

**Proposition 3.2.** *If $\mu_0$ is a premeasure on $\mathcal{A}$ and $\mu^*$ is given by* (3.4)*, then*

1. $\mu^*|_{\mathcal{A}} = \mu_0$;
2. *every set in $\mathcal{A}$ is $\mu^*$-measurable.*

*Proof.* To prove the first part, suppose $E \in \mathcal{A}$, $E \subseteq \bigcup_{j=1}^{\infty} A_j$ with $A_j \in \mathcal{A}$ and $B_n \doteq E \cap (A_n \setminus \bigcup_{j=1}^{n} A_j)$. Then, $B_n$'s are disjoint members of $\mathcal{A}$ whose union is $E \in \mathcal{A}$, hence by the definition of a premeasure and Lemma 3.1,

$$\mu_0(E) = \sum_{j=1}^{\infty} \mu_0(B_j) \leq \sum_{j=1}^{\infty} \mu_0(A_j);$$

taking the infimum with respect to $\{A_j\}_{j=1}^{\infty} \subseteq \mathcal{A}$ such that $E \subseteq \bigcup_{j=1}^{\infty} A_j$, we obtain $\mu_0(E) \leq \mu^*(E)$. On the other hand, taking $A_1 = E$ and $A_j = \varnothing$ for all $j \geq 2$, we have the reverse inequality $\mu^*(E) \leq \mu_0(E)$, meaning that $\mu^*(E) = \mu_0(E)$ for all $E \in \mathcal{A}$.

Next, let $E \subseteq X$ and $\varepsilon > 0$. Then, by the property (3.1) of the infimum, there exists $\{B_j\}_{j=1}^{\infty} \subseteq \mathcal{A}$ such that $E \subseteq \bigcup_{j=1}^{\infty} B_j$ and $\sum_{j=1}^{\infty} \mu_0(B_j) \le \mu^*(E) + \varepsilon$. Since $\mu_0$ is additive on $\mathcal{A}$, we have for any $A \in \mathcal{A}$

$$\mu^*(E) + \varepsilon \ge \sum_{j=1}^{\infty} \mu_0(B_j) = \sum_{j=1}^{\infty} \mu_0(B_j \cap A) + \sum_{j=1}^{\infty} \mu_0(B_j \cap A^c) \ge \mu^*(E \cap A) + \mu^*(E \cap A^c),$$

where the last inequality holds since $\bigcup_{j=1}^{\infty} B_j \cap A \supseteq E \cap A$ and $\bigcup_{j=1}^{\infty} B_j \cap A^c \supseteq E \cap A^c$. Since $\varepsilon$ is arbitrary, limiting $\varepsilon \to 0$ proves that $A$ is $\mu^*$-measurable. □

**Theorem 3.2.** *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra, $\mu_0$ a premeasure on $\mathcal{A}$, and $\mathcal{M} \doteq \sigma(\mathcal{A})$.*

1. *There exists a measure $\mu$ on $\mathcal{M}$ whose restriction is $\mu_0$, namely, $\mu = \mu^*|_{\mathcal{M}}$, where $\mu^*$ is given by* (3.4).
2. *If $\nu$ is another measure that extends $\mu_0$, then $\nu(E) \le \mu(E)$ for all $E \in \mathcal{M}$, with equality when $\mu(E) < \infty$.*
3. *If $\mu_0$ is $\sigma$-finite, then $\mu$ is the unique extension of $\mu_0$ to a measure on $\mathcal{M}$.*

*Proof.* Let $\mathcal{M}^*$ be the collection of all $\mu^*$-measurable sets. Then, by Carathéodory's Theorem, $\mathcal{M}^*$ is a $\sigma$-algebra and $\mu^*|_{\mathcal{M}^*}$ is a (complete) measure on $X$. By Proposition 3.2, $\mathcal{A} \subseteq \mathcal{M}^*$ and hence, $\mathcal{M} \subseteq \mathcal{M}^*$ by Lemma 1.1. Therefore, for any $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$ disjoint,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \mu^*|_{\mathcal{M}^*}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu^*|_{\mathcal{M}^*}(A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

Moreover, $\mu(\varnothing) = \mu^*(\varnothing) = 0$ is obvious by the definition of an outer measure. Therefore, $\mu = \mu^*|_{\mathcal{M}}$ is a measure on $\mathcal{M}$. Furthermore, $\mu|_{\mathcal{A}} = \mu_0$ by Proposition 3.2.

Next, suppose there exists another measure $\nu$ on $\mathcal{M}$ such that $\nu|_{\mathcal{A}} = \mu_0$. Then, $\mu(A) = \nu(A) = \mu_0(A)$ for any $A \in \mathcal{A}$. Hence, for any $E \in \mathcal{M}$ and $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{A}$ such that $E \subseteq \bigcup_{i=1}^{\infty} A_i$,

$$\nu(E) \le \sum_{i=1}^{\infty} \nu(A_i) = \sum_{i=1}^{\infty} \mu(A_i)$$

by monotonicity of a measure $\nu$ (Theorem 2.1a); taking the infimum for all such sets $\{A_i \in \mathcal{A}\}_{i=1}^{\infty}$ yields $\nu(E) \le \mu(E)$. Moreover, by Corollary 2.1 and $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$ for any $n$,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} \mu\left(\bigcup_{i=1}^{n} A_i\right) = \lim_{n\to\infty} \nu\left(\bigcup_{i=1}^{n} A_i\right) = \nu\left(\bigcup_{i=1}^{\infty} A_i\right). \tag{3.5}$$

Now, suppose $\mu(E) < \infty$ for $E \in \mathcal{M}$. Then, by the property (3.1) of the infimum and the definition (3.4), for each $\varepsilon > 0$, there exists $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{A}$ such that $E \subseteq \bigcup_{i=1}^{\infty} A_i$ and $\sum_{i=1}^{\infty} \mu_0(A_i) \le \mu(E) + \varepsilon$; since $\mu_0(A_i) = \mu(A_i)$ for all $A_i \in \mathcal{A}$ and $\mu$ is a measure on $\mathcal{M}$ $(= \sigma(\mathcal{A}))$, we have $\mu(A) \le \mu(E) + \varepsilon$, where $A \doteq \bigcup_{i=1}^{\infty} A_i$. Hence, $\mu(A \setminus E) \le \varepsilon$ by subtractivity of a measure $\mu$ (Theorem 2.1a). Finally, we establish

$$\mu(E) \le \mu(A) = \nu(A) = \nu(A) + \nu(A \setminus E) \le \nu(E) + \mu(A \setminus E) \le \nu(E) + \varepsilon.$$

To each (in)equality from left to right, we applied monotonicity (Theorem 2.1a), (3.5), subadditivity (Theorem 2.1a), $\nu(F) \le \mu(F)$ for $F = A \setminus E \in \mathcal{M}$, and $\mu(A \setminus E) \le \varepsilon$. Since $\varepsilon$ is arbitrary, limiting $\varepsilon \to 0$ proves $\mu(E) \le \nu(E)$ for all $E \in \mathcal{M}$ such that $\mu(E) < \infty$. As we already have $\nu(E) \le \mu(E)$ for any $E \in \mathcal{M}$, it is now obvious that $\mu(E) = \nu(E)$ for $E \in \mathcal{M}$ such that $\mu(E) < \infty$.

Finally, suppose $\mu_0$ is $\sigma$-finite, so that $X = \bigcup_{i=1}^{\infty} A_i$ for $\{A_i \in \mathcal{A}\}_{i=1}^{\infty}$ such that $\mu_0(A_i) < \infty$ for each $i$. Here, we assume that $\{A_i\}$ is disjoint without loss of generality.[1] Since $\mu(A_i) = \nu(A_i) = \mu_0(A_i) < \infty$, we have by monotonicity (Theorem 2.1a) $\mu(A_i \cap E) \le \mu(A_i) < \infty$ and $\nu(A_i \cap E) \le \nu(A_i) < \infty$ and thus $\mu(A_i \cap E) = \nu(A_i \cap E)$, for any $E \in \mathcal{M}$. Therefore, for any $E \in \mathcal{M}$,

$$\mu(E) = \sum_{i=1}^{\infty} \mu(E \cap A_i) = \sum_{i=1}^{\infty} \nu(E \cap A_i) = \nu(E),$$

that is, $\mu = \nu$ — $\mu$ is the unique extension of $\mu_0$ to a measure on $\mathcal{M}$. □

[1] otherwise, proceed with $\{B_i\}$ in place of $\{A_i\}$, where $B_1 \doteq A_1$ and $B_i \doteq A_i \setminus \bigcup_{j=1}^{i-1} A_j$ for all $i$. Obviously, $B_i \in \mathcal{A}$ for each $i$, $X = \bigcup_{i=1}^{\infty} B_i$, and $\{B_i\}_{i=1}^{\infty}$ is disjoint.

Note that in Theorem 3.2, generally $\mathcal{M} \subseteq \mathcal{M}^*$, and $\mu$ is not complete unless $\mathcal{M} = \mathcal{M}^*$, where $\mathcal{M}^*$ is the $\sigma$-algebra of all $\mu^*$-measurable sets as in the proof of Theorem 3.2. Moreover, the measure $\mu$ constructed by Theorem 3.2 is $\sigma$-finite if so is $\mu_0$, so that there exists $\{E_i \in \mathcal{A}\}_{i=1}^{\infty}$ such that $X = \bigcup_{i=1}^{\infty} E_i$ and $\mu_0(E_i) < \infty$ for all $i$ — obviously, $\mu(E_i) = \mu_0(E_i) < \infty$ by Proposition 3.4 and the $\sigma$-finiteness of $\mu$ follows. The next theorem provides a way to construct the completion of this $\sigma$-finite measure $\mu$ (and any $\sigma$-finite measure) using the outer measure similar to (3.4) and in a similar manner to the unique construction of the $\sigma$-finite measure $\mu$ from a $\sigma$-finite premeasure described above.

**Theorem 3.3.** *Let $(X, \mathcal{M}, \mu)$ be a $\sigma$-finite measure space, $\mu^*$ the outer measure induced by $\mu$ according to* (3.4)*, that is,*

$$\mu^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu(A_j) : A_j \in \mathcal{M},\ E \subseteq \bigcup_{j=1}^{\infty} A_j \right\} \text{ for any } E \subseteq X, \tag{3.6}$$

*and $\mathcal{M}^*$ the $\sigma$-algebra of $\mu^*$-measurable sets. Then, $\bar{\mu} = \mu^*|_{\mathcal{M}^*}$ is the completion of $\mu$.*

## 3.2 Borel and Lebesgue-Stieltjes Measures on $\mathbb{R}$

## 3.3 Measures and Integrations over a Product Space

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be $\sigma$-finite measure spaces. Recall that a (measurable) rectangle is a set of the form $M \times N$ for $M \in \mathcal{M}$ and $N \in \mathcal{N}$; by Corollary 1.5, the product $\sigma$-algebra $\mathcal{M} \otimes \mathcal{N}$ on $X \times Y$ corresponds to the $\sigma$-algebra generated by the set of all rectangles, i.e.,

$$\mathcal{M} \otimes \mathcal{N} = \sigma\big(\{M \times N : M \in \mathcal{M},\ N \in \mathcal{N}\}\big). \tag{3.7}$$

Let $\mathcal{A}$ be the collection of finite disjoint unions of rectangles. Then, by Corollary 1.5 again, $\mathcal{A}$ is an algebra, and the $\sigma$-algebra it generates corresponds to $\mathcal{M} \otimes \mathcal{N}$, i.e., $\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{A})$.

In this section, we construct a measure, called the product measure denoted by $\mu \times \nu$ on the product space $(X \times Y, \mathcal{M} \otimes \mathcal{N})$, with the special property:

$$\mu \times \nu\,(M \times N) = \mu(M) \cdot \nu(N) \text{ for all rectangles } M \times N. \tag{3.8}$$

Suppose $M \times N$ is a rectangle that is a finite or countable disjoint union of the rectangles $M_i \times N_i$. Then, for $x \in X$ and $y \in Y$, the indicator functions satisfy:

$$\mathbf{1}_M(x) \cdot \mathbf{1}_N(y) = \mathbf{1}_{M \times N}(x, y) = \textstyle\sum_i \mathbf{1}_{M_i \times N_i}(x, y) = \sum_i \mathbf{1}_{M_i}(x) \cdot \mathbf{1}_{N_i}(y)$$

Integrating $\mathbf{1}_M(x) \cdot \mathbf{1}_N(y)$ with respect to $\mu(x)$ and use the linearity of the integral, we obtain

$$\mu(M) \cdot \mathbf{1}_N(y) = \int \mathbf{1}_M(x) \cdot \mathbf{1}_N(y)\, d\mu(x) = \sum_i \int \mathbf{1}_{M_i}(x) \cdot \mathbf{1}_{N_i}(y)\, d\mu(x) = \sum_i \mu(M_i) \cdot \mathbf{1}_{N_i}(y);$$

integrating it with respect to $\nu(y)$ and using linearity again, we have

$$\begin{aligned} \mu(M) \cdot \nu(N) &= \int \mu(M) \cdot \mathbf{1}_N(y)\, d\nu(y) = \textstyle\sum_i \mu(M_i) \displaystyle\int \mathbf{1}_{N_i}(y)\, d\nu(y) \\ &= \textstyle\sum_i \mu(M_i) \cdot \nu(N_i). \end{aligned} \tag{3.9}$$

Since the algebra $\mathcal{A}$ is the set of all disjoint unions of rectangles, for each $E \in \mathcal{A}$, there exists a finite collection of disjoint rectangles $\{M_i \times N_i : M_i \in \mathcal{M},\ N_i \in \mathcal{N}\}_{i=1}^{n}$ such that $E = \bigcup_{i=1}^{n} M_i \times N_i$. With the convention $0 \cdot \infty = 0$, set

$$\pi(E) \doteq \sum_{i=1}^{n} \mu(M_i)\nu(N_i). \tag{3.10}$$

**Proposition 3.3.** *$\pi$ given by* (3.10) *is well-defined. That is, if* $\{A_j \times B_j : A_j \in \mathcal{M}.\ B_j \in \mathcal{N}\}_{j=1}^m$ *is another sequence of disjoint rectangles on* $X \times Y$ *such that* $E = \bigcup_{j=1}^m A_j \times B_j$*, then*

$$\pi(E) = \sum_{j=1}^m \mu(A_j) \cdot \nu(B_j) = \sum_{i=1}^n \mu(E_i) \cdot \nu(F_i).$$

*Moreover, $\pi$ is a premeasure on $\mathcal{A}$.*

*Proof.* Since $M_i \times N_i$ and $A_j \times B_j$ are rectangles and satisfy

$$M_i \times N_i = \bigcup_{j=1}^m (M_i \times N_i) \cap (A_j \times B_j) = \bigcup_{j=1}^m (M_i \cap A_j) \times (N_i \cap B_j)$$
$$A_j \times B_j = \bigcup_{i=1}^n (M_i \times N_i) \cap (A_j \times B_j) = \bigcup_{i=1}^n (M_i \cap A_j) \times (N_i \cap B_j)$$

where $\{(M_i \cap A_j) \times (N_i \cap B_j)\}_{i,j}$ is disjoint ($\because$ so are $\{M_i \times N_i\}_i$ and $\{A_j \times B_j\}_j$), we obtain by (3.9)

$$\pi(E) = \sum_{j=1}^m \mu(A_j) \cdot \nu(B_j) = \sum_{j=1}^m \sum_{i=1}^n \mu(M_i \cap A_j) \cdot \nu(N_i \cap B_j)$$
$$\pi(E) = \sum_{i=1}^n \mu(E_i) \cdot \nu(F_i) = \sum_{i=1}^n \sum_{j=1}^m \mu(M_i \cap A_j) \cdot \nu(N_i \cap B_j) = \sum_{j=1}^m \sum_{i=1}^n \mu(M_i \cap A_j) \cdot \nu(N_i \cap B_j),$$

and hence $\pi$ is well-defined.

Next, as $\varnothing = \varnothing \times \varnothing$ is a rectangle ($\varnothing \in \mathcal{M},\ \varnothing \in \mathcal{N}$), it is obvious that $\pi(\varnothing) = \mu(\varnothing) \cdot \nu(\varnothing) = 0$. To show the countable additivity, suppose that $\{E_i \in \mathcal{A}\}_{i=1}^\infty$ is disjoint and satisfies $E \doteq \bigcup_{i=1}^\infty E_i \in \mathcal{A}$. Then, for each $i$, there exists a sequence of disjoint rectangles $\{M_{i,j} \times N_{i,j} : M_{i,j} \in \mathcal{M},\ N_{i,j} \in \mathcal{N}\}_{j=1}^{n_i}$ such that $E_i = \bigcup_{j=1}^{n_i} N_{i,j} \times M_{i,j}$, hence $\pi(E_i) = \sum_{j=1}^{n_i} \mu(N_{i,j}) \cdot \nu(M_{i,j})$. Moreover, by $E \in \mathcal{A}$, there is a sequence of disjoint rectangles $\{A_k \times B_k : A_k \in \mathcal{M},\ B_k \in \mathcal{N}\}_{k=1}^m$ such that $E = \bigcup_{k=1}^m A_k \times B_k$, hence $\pi(E) = \sum_{k=1}^m \mu(A_k) \cdot \nu(B_k)$. Since the rectangles $A_k \times B_k$ and $M_{i,j} \times N_{i,j}$ satisfy

$$A_k \times B_k = (A_k \times B_k) \cap E = \bigcup_{i=1}^\infty \bigcup_{j=1}^{n_i} (A_k \times B_k) \cap (M_{i,j} \times N_{i,j}) = \bigcup_{i=1}^\infty \bigcup_{j=1}^{n_i} (A_k \cap M_{i,j}) \times (B_k \cap N_{i,j})$$
$$M_{i,j} \times N_{i,j} = E \cap (M_{i,j} \times N_{i,j}) = \bigcup_{k=1}^m (A_k \times B_k) \cap (M_{i,j} \times N_{i,j}) = \bigcup_{k=1}^m (A_k \cap M_{i,j}) \times (B_k \cap N_{i,j}),$$

where $\{(A_k \cap M_{i,j}) \times (B_k \cap N_{i,j})\}_{i,j,k}$ is a disjoint sequence of rectangles, we finally obtain by (3.9) that

$$\begin{aligned}\pi(E) &= \sum_{k=1}^m \mu(A_k) \cdot \nu(B_k) = \sum_{k=1}^m \left( \sum_{i=1}^\infty \sum_{j=1}^{n_i} \mu(A_k \cap M_{i,j}) \cdot \nu(B_k \cap N_{i,j}) \right) \\ &= \sum_{i=1}^\infty \sum_{j=1}^{n_i} \left( \sum_{k=1}^m \mu(A_k \cap M_{i,j}) \cdot \nu(B_k \cap N_{i,j}) \right) = \sum_{i=1}^\infty \sum_{j=1}^{n_i} \mu(M_{i,j}) \cdot \nu(N_{i,j}) = \sum_{i=1}^\infty \pi(E_i),\end{aligned}$$

hence $\pi$ is countably additive and thereby a premeasure on $\mathcal{A}$. □

By Proposition 3.3 and Theorem 3.2, the premeasure $\pi$ generates an outer measure $\pi^*$:

$$\pi^*(E) = \inf \left\{ \sum_{i=1}^\infty \pi(A_i) : A_i \in \mathcal{A},\ E \subseteq \bigcup_{i=1}^\infty A_i \right\} \quad \text{for any } E \subseteq X \times Y,$$

whose restriction to the product $\sigma$-algebra $\mathcal{M} \otimes \mathcal{N}$ is a measure that extends $\pi$. We call this measure the product of $\mu$ and $\nu$ and denote it by $\mu \times \nu$. By its construction, the product measure $\mu \times \nu$ satisfies (3.8).

Moreover, since $\mu$ and $\nu$ are $\sigma$-finite, there exist $\{A_i \in \mathcal{M}\}_{i=1}^{\infty}$ and $\{B_j \in \mathcal{N}\}_{j=1}^{\infty}$ such that $X = \bigcup_{i=1}^{\infty} A_i$ and $Y = \bigcup_{j=1}^{\infty} B_j$ with $\mu(A_i) < \infty$ and $\nu(B_j) < \infty$ for all $i$, $j$ — then $X \times Y = \bigcup_{i,j} A_i \times B_j$ and since $A_i \times B_j$ is a rectangle, $\mu \times \nu(A_i \times B_j) = \mu(A_i) \cdot \nu(B_j) < \infty$, meaning that the product $\mu \times \nu$ is also $\sigma$-finite; by Theorem 3.2, $\mu \times \nu$ is the unique extension on $\mathcal{M} \otimes \mathcal{N}$ that satisfies (3.8).

Now, we consider the $x$-section $E_x$ and $y$-section $E^y$ of $E \subseteq X \times Y$ defined as

$$E_x \doteq \{y \in Y : (x,y) \in E\} \text{ and } E^y \doteq \{x \in X : (x,y) \in E\}.$$

Similarly, we define the $x$-section $f_x$ and $y$-section $f^y$ of a function $f$ on $X \times Y$ as

$$f_x(y) = f^y(x) = f(x,y) \quad \forall (x,y) \in X \times Y.$$

Note that:

1. if $E \in \mathcal{M} \otimes \mathcal{N}$, then $E_x \in \mathcal{N}$ for all $x \in X$ and $E^y \in \mathcal{M}$ for all $y \in Y$ (Proposition 1.11);
2. if $E, F \in \mathcal{M} \otimes \mathcal{N}$, then $(E \cap F)_x = E_x \cap F_x$ and $(E \cap F)^y = E^y \cap F^y$ since, e.g., for the former,

$$\begin{aligned}(E \cap F)_x &= \{y \in Y : (x,y) \in E \cap F\} \\ &= \{y \in Y : (x,y) \in E \text{ and } (x,y) \in F\} \\ &= \{y \in Y : (x,y) \in E\} \cap \{y \in Y : (x,y) \in F\} = E_x \cap F_y.\end{aligned}$$

   Similarly, $(E \cup F)_x = E_x \cup F_x$ and $(E \cup F)^y = E^y \cup F^y$ (note: $(x,y) \in E \cup F \iff (x,y) \in E$ *or* $\in F$).
3. For an indicator function $\mathbf{1}_E$ ($E \in \mathcal{M} \otimes \mathcal{N}$), $(\mathbf{1}_E)_x = \mathbf{1}_{E_x}$ and $(\mathbf{1}_E)^y = \mathbf{1}_{E^y}$
4. if $f$ is $(\mathcal{M} \otimes \mathcal{N}, \mathcal{O})$-measurable, then $f_x$ is $(\mathcal{N}, \mathcal{O})$-measurable for all $x \in X$ and $f^y$ is $(\mathcal{M}, \mathcal{O})$-measurable for all $y \in Y$ (Proposition 1.12).

**Theorem 3.4.** *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, and*

$$\mu \times \nu(E) = \int \nu(E_x)\, d\mu(x) = \int \mu(E^y)\, d\nu(y). \tag{3.11}$$

*Proof. (when $\mu$ and $\nu$ are finite)* Let $\mathcal{C} = \{E \in \mathcal{M} \otimes \mathcal{N} : \text{ the conclusions of the theorem hold on } E\}$. Then, by its construction, we have $\mathcal{C} \subseteq \mathcal{M} \otimes \mathcal{N}$, hence the proof is complete if $\mathcal{M} \otimes \mathcal{N} \subseteq \mathcal{C}$. First, we show that $\{M \times N : M \in \mathcal{M},\ N \in \mathcal{N}\} \subseteq \mathcal{C}$. Let $E = M \times N$ for $M \in \mathcal{M}$ and $N \in \mathcal{N}$. Then, since

$$E_x = \begin{cases} N \text{ for } x \in M \\ \varnothing \text{ for } x \notin M \end{cases} \quad \text{and} \quad E^y = \begin{cases} \varnothing \text{ for } y \notin N \\ M \text{ for } y \in N, \end{cases}$$

we have $\nu(E_x) = \nu(N) \cdot \mathbf{1}_M(x)$ and $\mu(E^y) = \mu(M) \cdot \mathbf{1}_N(y)$, hence the maps $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable by Proposition 1.25 (note that $\nu(N) < \infty$ and $\mu(M) < \infty$ as the measures are finite). Moreover, by Theorem 2.4,

$$\int \nu(E_x)\, d\mu(x) = \nu(N) \cdot \int \mathbf{1}_M\, d\mu = \nu(N) \cdot \mu(M), \quad \text{and} \int \mu(E^y)\, d\nu(y) = \mu(M) \cdot \int \mathbf{1}_N\, d\nu = \mu(M) \cdot \nu(N),$$

where $\mu(M) \cdot \nu(N) = \mu \times \nu(M \times N)$ ($\because M \in \mathcal{M}$ and $N \in \mathcal{N}$). Therefore, $E \in \mathcal{C}$, meaning that $\{M \times N : M \in \mathcal{M},\ N \in \mathcal{N}\} \subseteq \mathcal{C}$. Moreover, the collection $\mathcal{A}$ of disjoint unions of rectangles is an algebra by Corollary 1.5. Let $\{E_i\}_{i=1}^n$ be a disjoint sequence of rectangles such that $E = \bigcup_{i=1}^n E_i$ for $E \in \mathcal{A}$ and denote $E_i = M_i \times N_i$ for some $M_i \in \mathcal{M}$ and $N_i \in \mathcal{N}$. If $x \in M_i \cap M_j$, then $(E_i)_x = N_i$ and $(E_j)_x = N_j$, where $N_i$ and $N_j$ *must* be disjoint as $E_i$ and $E_j$ are disjoint. Moreover, it is obvious that $(E_i)_x = \varnothing$ whenever $x \notin M_i$, hence the set $\{(E_i)_x\}_{i=1}^n$ is disjoint. Since $E_x = \bigcup_{i=1}^n (E_i)_x$ (see the proof of Proposition 1.11), we finally obtain

$$\nu(E_x) = \sum_{i=1}^n \nu((E_i)_x) \text{ and likewise, } \mu(E^y) = \sum_{i=1}^n \mu((E_i)^y), \tag{3.12}$$

where $E_i$'s are all rectangles. Therefore, the maps $x \mapsto \mu(E_x)$ and $y \mapsto \mu(E^y)$ are also measurable *for any* $E \in \mathcal{A}$. Moreover, as (3.12) implies $\nu(E_x) = \sum_{i=1}^n \nu(N_i) \cdot \mathbf{1}_{M_i}(x)$ and $\mu(E^y) = \sum_{i=1}^n \mu(M_i) \cdot \mathbf{1}_{N_i}(y)$, we obtain:

$$\mu \times \nu(E) = \sum_{i=1}^n \mu \times \nu(E_i) = \begin{cases} \displaystyle\sum_{i=1}^n \int \nu(N_i) \cdot \mathbf{1}_{M_i}(x)\, d\mu(x) = \int \nu(E_x)\, d\mu(x) \\ \displaystyle\sum_{i=1}^n \int \mu(M_i) \cdot \mathbf{1}_{N_i}(y)\, d\nu(y) = \int \mu(E^y)\, d\nu(y), \end{cases}$$

from which we conclude that $E \in \mathcal{C}$, that is, $\mathcal{A} \subseteq \mathcal{C}$.

The remaining process is to show that $\mathcal{C}$ is a monotone class. Let $\{E_n\}_{n=1}^\infty$ be an increasing sequence in $\mathcal{C}$ and $E = \bigcup_{n=1}^\infty E_n$. Then, $f_n(y) := \mu((E_n)^y)$ is measurable by the definition of $\mathcal{C}$ and pointwise increases to $f(y) := \mu(E^y)$ by Theorem 2.1a and 2.1c. Hence, $f$ is measurable by Proposition 1.26 and by the MCT (Theorem 2.3),

$$\int \mu(E^y)\, d\nu(y) = \int f\, d\nu = \lim_{n\to\infty} \int f_n\, d\nu = \lim_{n\to\infty} \int \mu((E_n)^y)\, d\nu(y) = \lim_{n\to\infty} \mu \times \nu(E_n) = \mu \times \nu(E),$$

where the last equality comes from the application of Theorem 2.1c for the measure $\mu \times \nu$. Likewise, we also have $\mu \times \nu(E) = \int \nu(E_x)\, d\mu(x)$, hence $E \in \mathcal{C}$. Similarly, if $\{E_n\}_{n=1}^\infty$ is a decreasing sequence in $\mathcal{C}$ and $E = \bigcap_{n=1}^\infty E_n$, then $g_n(y) := \mu((E_n)^y)$ for each $n$ is measurable by the definition of $\mathcal{C}$ and, since $\mu((E_n)^y) \le \mu(X)$, uniformly bounded (by $\mu(X) < \infty$) for all $n$. Moreover, since $\mu((E_1)^y) \le \mu(X) < \infty$, $g_n(y)$ pointwise decreases to $g(y) := \mu((E_n)^y)$ by Theorem 2.1a and 2.1d. Therefore, by the bounded convergence theorem (Corollary 2.11) and $\nu(Y) < \infty$, we have $g \in \mathcal{L}^1(\nu)$ and $\lim_{n\to\infty} \int g_n\, d\nu = \int g\, d\nu$, that is, $\mu \times \nu(E) = \int \mu(E^y)\, d\nu(y)$ as above and likewise, $\mu \times \nu(E) = \int \nu(E_x)\, d\mu(x)$. Therefore, $E \in \mathcal{C}$ and thus $\mathcal{C}$ is a monotone class.

Since $\mathcal{C}$ is a monotone class, $\mathcal{A} \subseteq \mathcal{C}$, and $\sigma(\mathcal{A}) = \mathcal{M} \otimes \mathcal{N}$, applying the monotone class theorem (Theorem B.1) and Lemma B.1 results in $\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{A}) = \mathfrak{c}(\mathcal{A}) \subseteq \mathcal{C}$.

*(when $\mu$ and $\nu$ are $\sigma$-finite)* In this case, we can write $X \times Y$ as a union of an increasing sequence $\{X_i \times Y_i\}_{i=1}^\infty$ of rectangles, that is,

$$X \times Y = \bigcup_{i=1}^\infty X_i \times Y_i \quad \text{with} \quad X_i \times Y_i \subseteq X_{i+1} \times Y_{i+1}, \quad X_i \in \mathcal{M}, \quad Y_i \in \mathcal{N}, \text{ for all } i \in \mathbb{N},$$

which implies that $X_i \subseteq X_{i+1}$ and $Y_i \subseteq Y_{i+1}$ for all $i$ as well as $X = \bigcup_{i=1}^\infty X_i$ and $Y = \bigcup_{i=1}^\infty Y_i$.

For $E \in \mathcal{M} \otimes \mathcal{N}$, since

$$\begin{aligned}(E \cap (X_i \times Y_i))_x &= E_x \cap (X_i \times Y_i)_x = (X \times E_x)_x \cap (X_i \times Y_i)_x = ((X \times E_x) \cap (X_i \times Y_i))_x \\ &= ((X_i \times (E_x \cap Y_i))_x,\end{aligned}$$

applying the preceding argument to $E \cap (X_j \times Y_j)$ results in

$$\begin{aligned}\mu \times \nu(E \cap (X_i \times Y_i)) &= \int \nu((E \cap (X_i \times Y_i))_x)\, d\mu(x) = \int \nu((X_i \times (E_x \cap Y_i))_x)\, d\mu(x) \\ &= \int \mathbf{1}_{X_i}(x) \cdot \nu(E_x \cap Y_i))\, d\mu(x).\end{aligned}$$

Therefore, by the monotone convergence theorem (Theorem 2.3) and Theorem 2.1c, we obtain

$$\mu(E) = \mu(E \cap (X \times Y)) = \lim_{i\to\infty} \mu \times \nu(E \cap (X_i \times Y_i)) = \int \lim_{i\to\infty} \mathbf{1}_{X_i}(x) \cdot \nu(E_x \cap Y_i)\, d\mu(x) = \int \nu(E_x)\, d\mu(x).$$

Likewise, $\mu \times \nu(E) = \int \mu(E^y)\, d\nu(y)$. Here, the measurability of $x \mapsto \nu(E_x)$ and $y \mapsto \nu(E^y)$ is ensured by Proposition 1.26. $\square$

**Theorem 3.5** (The Fubini-Tonelli Theorem)**.** *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are $\sigma$-finite measure spaces.*

1. *(Tonelli) If $f \in \mathcal{L}^+(X \times Y)$, then the functions $g(x) = \int f_x d\,\nu$ and $h(x) = \int f^y \, d\mu$ belong to $\mathcal{L}^+(X)$ and $\mathcal{L}^+(Y)$, respectively, and*

$$\int f \, d(\mu \times \nu) = \int \left[ \int f(x,y) \, d\nu(y) \right] d\mu(x) = \int \left[ \int f(x,y) \, d\mu(x) \right] d\nu(y) \tag{3.13}$$

2. *(Fubini) If $f \in \mathcal{L}^1(X \times Y)$, then $f_x \in \mathcal{L}^1(\nu)$ for a.e. $x \in X$, $f^y \in \mathcal{L}^1(\mu)$ for a.e. $y \in Y$, the a.e.-defined functions $g(x) = \int f_x d\,\nu$ and $h(x) = \int f^y \, d\mu$ are in $\mathcal{L}^1(\mu)$ and $\mathcal{L}^1(\nu)$, respectively, and* (3.13) *holds.*

*Proof.* The Tonelli's theorem reduces to Theorem 3.4 in case $f$ is an indicator function and thus by the linearity of the integral (e.g., Proposition 2.5), it holds for nonnegative simple functions. For example, if $f = \mathbf{1}_E$ for $E \in \mathcal{M} \otimes \mathcal{N}$, then

$$g(x) = \int (\mathbf{1}_E)_x \, d\nu = \int \mathbf{1}_{E_x} \, d\nu = \nu(E_x), \quad h(y) = \int (\mathbf{1}_E)^y \, d\mu = \int \mathbf{1}_{E^y} \, d\mu = \mu(E^y),$$

and by Theorem 3.4, $\int \mathbf{1}_E(x,y) = \begin{cases} \int \nu(E_x) \, d\mu(x) = \int \left[ \int \mathbf{1}_E(x,y) \, d\nu(y) \right] d\mu(x), \\ \int \nu(E_x) \, d\mu(x) = \int \left[ \int \mathbf{1}_E(x,y) \, d\nu(y) \right] d\mu(x). \end{cases}$

If $f \in \mathcal{L}^+(X \times Y)$, then there exists a sequence $\{f_n\}$ of nonnegative simple functions that increases pointwise to $f$ by Theorem 1.1. Then, the MCT (Theorem 2.3) implies:

1. the corresponding $g_n$ and $h_n$ pointwise increase $g$ and $h$ as shown below: by Proposition 2.5,

$$g_n = \int (f_n)_x \, d\nu \le \int (f_{n+1})_x \, d\nu = g_{n+1} \text{ and thus } \lim_{n\to\infty} g_n = \int \lim_{n\to\infty} (f_n)_x \, d\nu = \int f_x \, d\nu = g,$$
$$h_n = \int (f_n)^y \, d\mu \le \int (f_{n+1})^y \, d\mu = g_{n+1} \text{ and thus } \lim_{n\to\infty} h_n = \int \lim_{n\to\infty} (f_n)^y \, d\nu = \int f^y \, d\mu = h$$

(so that $g$ and $h$ are measurable by Proposition 1.26);

2. (3.13) holds as shown below:

$$\int g \, d\mu = \lim_{n\to\infty} \int g_n \, d\mu = \lim_{n\to\infty} \int f_n \, d(\mu \times \nu) = \int f \, d(\mu \times \nu),$$
$$\int h \, d\nu = \lim_{n\to\infty} \int h_n \, d\nu = \lim_{n\to\infty} \int f_n \, d(\mu \times \nu) = \int f \, d(\mu \times \nu).$$

This proves the Tonelli's Theorem and also shows that if $f \in \mathcal{L}^+(X \times Y)$ and $\int f \, d(\mu \times \nu) < \infty$, then $g < \infty$ and $h < \infty$ a.e. by Proposition 2.10. That is, $f_x \in \mathcal{L}^1(\nu)$ for a.e. $x$ and $f_y \in \mathcal{L}^1(\mu)$ for a.e. $y$. On the other hand, if $f \in \mathcal{L}^1(\mu \times \nu)$, then it is equivalent to the existence of $f_+, f_- \in \mathcal{L}^+(X \times Y)$ s.t. $f = f_+ - f_-$ and $\int f_+ \, d(\mu \times \nu) < \infty$ and $\int f_- \, d(\mu \times \nu) < \infty$ by Definition 2.1 and Corollary 1.11. Therefore, the proof of Fubini's theorem can be done by applying the above results to $f_+$ and $f_-$ and combine them using $f = f_+ - f_-$. □

**Remark 3.1.** *Even if $\mu$ and $\nu$ are complete, $\mu \times \nu$ is almost never complete. For example, suppose that there is a non-empty $A \in \mathcal{M}$ with $\mu(A) = 0$ and that $\mathcal{N} \neq \mathcal{P}(Y)$ (e.g., $\mu = \nu =$ the Lebesgue measure on $\mathbb{R}$). If $E = \mathcal{P} \setminus \mathcal{N}$, then $A \times E \notin \mathcal{M} \otimes \mathcal{N}$ by the contraposition of Proposition 1.11: for $E \subseteq X \times Y$, if $E_x \notin \mathcal{N}$ for some $x$ or $E^y \notin \mathcal{M}$ for some $y$, then $E \notin \mathcal{M} \otimes \mathcal{N}$. However, $A \times E \subset A \times Y$ and $\mu \times \nu(A \times Y) = 0$. For an extension of the Tonelli-Fubini's theorem to a completion of $\mu \times \nu$, see (Folland, 1999).*

## 3.4 Modes of Convergence

**Proposition 3.4.** *If $f_n \to f$ in $\mathcal{L}^1$, then $f_n \to f$ in measure.*

*Proof.* Given $\epsilon > 0$, let $E_{n,\epsilon} = \{x \in X : |f_n(x) - f(x)| \geq 0\} = |f_n - f|^{-1}([\epsilon, \infty))$. Then,

$$\infty > \int_X |f_n - f| \geq \int_{E_{n,\epsilon}} |f_n - f| \geq \epsilon \cdot \mu(E_{n,\epsilon}).$$

Therefore, $\mu(E_{n,\epsilon}) \leq \epsilon^{-1} \cdot \int_X |f_n - f| \to 0$ as $n \to \infty$. □

**Theorem 3.6.**

1. *If $\{f_n\}$ is Cauchy in measure, then there exists a measurable function $f$ such that $f_n \to f$ in measure.*
2. *If $f_n \to f$ in measure, then there exists a subsequence $\{f_{n_j}\}$ that converges to $f$ a.e.*
3. *If $f_n \to f$ and $f_n \to g$ both in measure, then $f = g$ a.e.*

Proof of Theorem 2.30 with Borel-Cantalli Lemma

## 3.5 $\mathcal{L}^p$ Spaces

## 3.6 Signed Measures and their Decompositions

A *signed measure* on a measurable space $(X, \mathcal{M})$ is a function $\nu : \mathcal{M} \to [-\infty, \infty]$ s.t.

1. $\nu(\varnothing) = 0$;
2. $\nu$ assumes at most one of the values $\pm\infty$, i.e., either $-\infty < \inf_{E \in \mathcal{M}} \nu(E)$ or $\sup_{E \in \mathcal{M}} \nu(E) < \infty$ must be true;
3. if $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{M}$ is disjoint, then $\nu\big(\bigcup_{i=1}^{\infty} E_i\big) = \sum_{i=1}^{\infty} \nu(E_i)$, where the latter sum converges absolutely if finite, that is,
$$\left|\sum_{i=1}^{\infty} \nu(E_i)\right| < \infty \implies \sum_{i=1}^{\infty} |\nu(E_i)| < \infty. \tag{3.14}$$

Note that the converse of (3.14) is also true as any absolutely convergent infinite sum of a real sequence always converges. Here, the absolute convergence condition (3.14) is required to well-define a signed measure $\nu$—if the sum is merely conditionally convergent (so that the precondition in (3.14) is still valid), then by Riemann's rearrangement theorem, there exist permutations $\sigma_1$ and $\sigma_2$ s.t. $\nu\big(\bigcup_{i=1}^{\infty} E_{\sigma_1(i)}\big) = \sum_{i=1}^{\infty} \nu(E_{\sigma_1(i)}) = \infty$ and $\nu\big(\bigcup_{i=1}^{\infty} E_{\sigma_2(i)}\big) = \sum_{i=1}^{\infty} \nu(E_{\sigma_2(i)}) = -\infty$ (already violating the second requirement of the definition), whereas

$$\nu\big(\textstyle\bigcup_{i=1}^{\infty} E_{\sigma_1(i)}\big) = \nu\big(\bigcup_{i=1}^{\infty} E_{\sigma_2(i)}\big) = \nu\big(\bigcup_{i=1}^{\infty} E_i\big) = \sum_{i=1}^{\infty} \nu(E_i),$$

where the last sum is finite by conditional convergence.

## 3.7 Lebesgue-Radon-Nikodym Theorem

# Chapter 4

# Lebesgue Measure, Integral, and Differentiation

## 4.1 Motivations

## 4.2 Lebesgue Measure and Integral

## 4.3 Lebesgue Differentiation on Euclidean Space

## 4.4 Functions of Bounded Variations

# Chapter 5

# Stochastic Processes

## 5.1 Fundamentals in Probability

the second Borel-Cantalli Lemma...

## 5.2 Itô Integrals

**Definition 5.1.** *Let* $\mathcal{V} \doteq \mathcal{V}(S,T)$ *be a class of functions*

$$f(t,w) : [0,\infty) \times \Omega \to \mathbb{R}$$

*such that*

1. $f$ *is* $\mathcal{B} \times \mathcal{F}$*-measurable, where* $\mathcal{B} \doteq \mathcal{B}_{[0,\infty)}$ *is the Borel* $\sigma$*-algebra on* $[0,\infty)$*;*
2. $f$ *is* $\mathcal{F}_t$*-adapted;*
3. $\mathbb{E}\big[\int_T^S f^2(t,w)\,dt\big] < \infty$.

**Lemma 5.1.** *Let* $g \in \mathcal{V}$ *be bounded and* $g(\cdot,w)$ *is continuous for each* $w \in \Omega$*. Then, there exists a sequence of elementary functions* $\{\phi_n \in \mathcal{V}\}$ *such that*

$$\mathbb{E}\left[\int_T^S (g(t,w) - \phi_n(t,w))^2\,dt\right] \to 0 \text{ as } n \to \infty.$$

*Proof.* □

# Bibliography

Folland, G. B. *Real analysis: modern techniques and their applications.* John Wiley & Sons, 1999.

# Appendix A

# Elementary Family

An elementary family, defined below, plays a role in constructing an algebra which is the domain of a premeasure and also used in the monotone class lemma in Appendix B.

**Definition A.1.** *A non-empty family $\mathcal{E} \subseteq \mathcal{P}(X)$ is said to be an elementary family of $X$ iff*

1. $\phi \in \mathcal{E}$;
2. *if* $E, F \in \mathcal{E}$, *then* $E \cap F \in \mathcal{E}$;
3. *if* $E \in \mathcal{E}$, *then* $E^c$ *is a finite disjoint union of members of* $\mathcal{E}$.

By the first and the third properties, $X$ is a finite disjoint union of members of an elementary family of $X$. In the following, we show that the family of finite disjoint union of an elementary family forms an algebra.

**Proposition A.1.** *If $\mathcal{E}$ is an elementary family of $X$, then the family $\mathcal{A}$ of all finite disjoint unions of members of $\mathcal{E}$ is an algebra on $X$.*

*Proof.* Let $A_1, \cdots, A_n \in \mathcal{E}$ and $A_n^c = \bigcup_{k=1}^m B_k$ for some disjoint $B_k \in \mathcal{E}$. Then, for $j \in \{1, 2, \cdots, n\}$,

$$A_j \setminus A_n = A_j \cap A_n^c = \bigcup_{k=1}^{m} (A_j \cap B_k) \in \mathcal{A} \quad (\because A_j \cap B_k \in \mathcal{E} \text{ and } \{A_j \cap B_k\}_{k=1}^m \text{ is disjoint.}).$$

This implies $\bigcup_{j=1}^n A_j = A_n \cup \left( \bigcup_{j=1}^{n-1} (A_j \setminus A_n) \right) \in \mathcal{A}$. To show that $\mathcal{A}$ is closed under complements, let $A_j^c = \bigcup_{k=1}^{m_j} B_{j,k}$ for $j \in \{1, 2, \cdots, n\}$, where $B_{j,1}, B_{j,2}, \cdots, B_{j,m_j}$ are disjoint members of $\mathcal{E}$. Then,

$$\left( \bigcup_{j=1}^{n} A_j \right)^c = \bigcap_{j=1}^{n} \left( \bigcup_{k=1}^{m_j} B_{j,k} \right) = \bigcup \Big\{ \underbrace{B_{1,k_1} \cap \cdots \cap B_{n,k_n}}_{\in \mathcal{E}} : 1 \le k_j \le m_j,\ 1 \le j \le n \Big\} \in \mathcal{A}$$

by the set operations over a finite number of sets, which completes the proof. □

# Appendix B

# The Montone Class Theorem

A technical concept, related to Tonelli-Fubini's Theorem, is a *monotone class* $\mathcal{C} \subseteq \mathcal{P}(X)$, which is a collection of subsets of $X$ that is closed under countable unions and countable decreasing intersections, i.e.,

$$\begin{aligned} &1.\ \{E_i\}_{i=1}^{\infty} \subseteq \mathcal{C} \text{ and } E_1 \subseteq E_2 \subseteq E_3 \cdots \implies \textstyle\bigcup_{i=1}^{\infty} E_i \in \mathcal{C}, \\ &2.\ \{E_i\}_{i=1}^{\infty} \subseteq \mathcal{C} \text{ and } E_1 \supseteq E_2 \supseteq E_3 \cdots \implies \textstyle\bigcap_{i=1}^{\infty} E_i \in \mathcal{C}. \end{aligned} \tag{B.1}$$

A $\sigma$-algebra is closed under countable unions and intersections by its definition and Proposition 1.1 and thus a monotone class, but not vice versa.

**Proposition B.1.** *Let $\{\mathcal{C}_\alpha\}_{\alpha \in A}$ be a family of monotone classes on $X$. Then, $\mathcal{C} = \bigcap_{\alpha \in A} \mathcal{C}_\alpha$ is a monotone class on $X$.*

*Proof.* Assume $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{C}$ and $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$. Then, $E_i \in \mathcal{C}$ implies that $E_i \in \mathcal{C}_\alpha$ for all $\alpha \in A$. Since $\mathcal{C}_\alpha$ is a monotone class and $\{E_i\}_{i=1}^{\infty}$ is increasing, we obtain $\bigcup_{i=1}^{\infty} E_i \in \mathcal{C}_\alpha$ for all $\alpha \in A$, implying that $\bigcup_{i=1}^{\infty} E_i \in \mathcal{C}$. Similarly, for $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{C}$ such that $E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots$, one can also show that $\bigcap_{i=1}^{\infty} E_i$ belongs to $\mathcal{C}_\alpha$ for all $\alpha \in A$ and thus $\mathcal{C}$. Hence, $\mathcal{C}$ is a monotone class. $\square$

**Definition B.1.** *For any family $\mathcal{E} \subseteq \mathcal{P}(X)$, $\mathfrak{c}(\mathcal{E})$ denotes the smallest monotone class on $X$ that contains $\mathcal{E}$; we call $\mathfrak{c}(\mathcal{E})$ the monotone class generated by $\mathcal{E}$.*

**Remark B.1.** *Since the largest $\sigma$-algebra $\mathcal{P}(X)$ is also the largest monotone class on $X$, there is at least one monotone class $\mathcal{P}(X)$ that contains $\mathcal{E}$. Moreover, since any (uncountable) intersection of monotone classes is also a monotone class as shown in Proposition B.1, the smallest monotone class $\mathfrak{c}(\mathcal{E})$ in Definition B.1 can be constructed and recognized as the intersection of all monotone classes containing $\mathcal{E}$, similarly to the $\sigma$-algebra cases.*

By Proposition B.1, the following lemma is obvious.

**Lemma B.1.** *If $\mathcal{E} \subseteq \mathcal{C}$, then $\mathfrak{c}(\mathcal{E}) \subseteq \mathcal{C}$.*

The next lemma shows that $\sigma$-algebra and monotone class both generated by an algebra coincide.

**Theorem B.1** (The Monotone Class Theorem)**.** *If $\mathcal{A}$ is an algebra on $X$, then $\sigma(\mathcal{A}) = \mathfrak{c}(\mathcal{A})$.*

*Proof.* Since any $\sigma$-algebra is a monotone class, $\sigma(\mathcal{A})$ is a monotone class and hence, $\sigma(\mathcal{A}) \supseteq \mathfrak{c}(\mathcal{A})$. Next, if we show that $\mathfrak{c}(\mathcal{A})$ is a $\sigma$-algebra on $X$, then we have $\sigma(\mathcal{A}) \subseteq \mathfrak{c}(\mathcal{A})$ and hence $\sigma(\mathcal{A}) = \mathfrak{c}(\mathcal{A})$. To show that $\mathfrak{c}(\mathcal{A})$ is a $\sigma$-algebra on $X$, first note that by $\phi, X \in \mathcal{A}$ (Proposition 1.1) and $\mathcal{A} \subseteq \mathfrak{c}(\mathcal{A})$, we have $\phi, X \in \mathfrak{c}(\mathcal{A})$. Next, for $E \in \mathfrak{c}(A)$, define

$$\mathcal{C}(E) \doteq \big\{F \in \mathfrak{c}(\mathcal{A}) : E \setminus F,\ F \setminus E, \text{ and } E \cap F \text{ are all in } \mathfrak{c}(\mathcal{A})\big\}.$$

Then, we see that $\phi, E \in \mathcal{C}(E)$ and that

$$\forall E, F \in \mathfrak{c}(\mathcal{A}) : \quad F \in \mathcal{C}(E) \iff E \in \mathcal{C}(F). \tag{B.2}$$

Moreover, for $\{F_j \in \mathcal{C}(E)\}_{i=1}^{\infty}$ such that $F_j \subseteq F_{j+1}\ \forall i \in \mathbb{N}$, we have $\bigcup_{i=1}^{\infty} F_j \in \mathfrak{c}(\mathcal{A})$ and furthermore, since $E \setminus F_j \supseteq E \setminus F_{j+1}$, $F_j \setminus E \subseteq F_{j+1} \setminus E$, and $E \cap F_j \subseteq E \cap F_{j+1}$ for all $i \in \mathbb{N}$ and all the sets belong to the monotone class $\mathfrak{c}(\mathcal{A})$, we have

$$\begin{aligned}
&E \setminus \left(\bigcup_{i=1}^{\infty} F_j\right) = E \cap \left(\bigcap_{i=1}^{\infty} F_j^c\right) = \bigcap_{i=1}^{\infty}(E \cap F_j^c) = \bigcap_{i=1}^{\infty} E \setminus F_j \in \mathfrak{c}(\mathcal{A}),\\
&\left(\bigcup_{i=1}^{\infty} F_j\right) \setminus E = \left(\bigcup_{i=1}^{\infty} F_j\right) \cap E^c = \bigcup_{i=1}^{\infty}(F_j \cap E^c) = \bigcup_{i=1}^{\infty} F_j \setminus E \in \mathfrak{c}(\mathcal{A}),\\
&E \cap \left(\bigcup_{i=1}^{\infty} F_j\right) = \bigcup_{i=1}^{\infty}(E \cap F_j) \in \mathfrak{c}(\mathcal{A}).
\end{aligned}$$

Hence, $\bigcup_{i=1}^{\infty} F_j \in \mathcal{C}(E)$. Similarly, one can show for a decreasing sequence $\{F_j \in \mathcal{C}(E)\}$ that $\bigcap_{i=1}^{\infty} F_j \in \mathcal{C}(E)$. Therefore, $\mathcal{C}(E)$ is a monotone class. Moreover, since $\mathcal{A}$ is an algebra, $E \in \mathcal{A}$ implies $F \in \mathcal{C}(E)$ for all $F \in \mathcal{A}$ by the definition of $\mathcal{C}(E)$ and Proposition 1.1. That is, $\mathcal{A} \subseteq \mathcal{C}(E)$ and by Lemma B.1, $\mathfrak{c}(\mathcal{A}) \subseteq \mathcal{C}(E)$ for any $E \in \mathcal{A}$. That is, for any $F \in \mathfrak{c}(\mathcal{A})$ and $E \in \mathcal{A}$, we have $F \in \mathcal{C}(E)$, which in turn implies $E \in \mathcal{C}(F)$ by (B.2). Therefore, we further obtain $\mathcal{A} \subseteq \mathcal{C}(F)$ for any $F \in \mathfrak{c}(\mathcal{A})$, and by Lemma B.1, $\mathfrak{c}(\mathcal{A}) \subseteq \mathcal{C}(F)$ for any $F \in \mathfrak{c}(\mathcal{A})$. Conclusions: if $E, F \in \mathfrak{c}(\mathcal{A})$, then both $E \setminus F$ and $E \cap F$ also belong to $\mathfrak{c}(\mathcal{A})$. Since $\phi, X \in \mathfrak{c}(\mathcal{A})$, we can prove that $\mathfrak{c}(\mathcal{A})$ is an algebra by checking that it is closed under complements and finite unions:

$$\begin{aligned}
&1)\ E \in \mathfrak{c}(\mathcal{A}) \implies E^c = X \setminus E \in \mathfrak{c}(\mathcal{A}),\\
&2)\ \{E_i\}_{j=1}^{N} \subseteq \mathfrak{c}(\mathcal{A}) \implies \bigcup_{j=1}^{N} E_i = \left(\bigcap_{j=1}^{N} E_i^c\right)^c \in \mathfrak{c}(\mathcal{A}),
\end{aligned}$$

where the second property is deduced from the first and $E \cap F \in \mathfrak{c}(\mathcal{A})$ for $E, F \in \mathfrak{c}(\mathcal{A})$. Therefore, for any $\{E_i\}_{j\in\mathbb{N}} \subseteq \mathfrak{c}(\mathcal{A})$, we have $\bigcup_{j=1}^{N} E_i \in \mathfrak{c}(\mathcal{A})$ for all $N \in \mathbb{N}$, where $\left\{\bigcup_{j=1}^{N} E_i\right\}_{N=1}^{\infty}$ is an increasing sequence in $\mathfrak{c}(\mathcal{A})$. Therefore, since $\mathfrak{c}(\mathcal{A})$ is also a monotone class, we have $\bigcup_{i=1}^{\infty} E_i \in \mathfrak{c}(\mathcal{A})$ by (B.1). That is, $\mathfrak{c}(\mathcal{A})$ is a $\sigma$-algebra and the proof is completed. $\square$