Measure Theory & Probability

Joe Erickson

Contents

1

Measure Spaces

1.1 – Sequences and Set Operations

The set of real numbers we denote by $\mathbb{R}$ or $(-\infty, \infty)$, and the set of **extended real numbers** we take to be the set

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\} = [-\infty, \infty].$$

Often $+\infty$ is denoted by ∞.

Given $x_1, x_2 \in \overline{\mathbb{R}}$, we define

$$x_1 \vee x_2 = \max\{x_1, x_2\} \quad \text{and} \quad x_1 \wedge x_2 = \min\{x_1, x_2\}.$$

More generally if $x_1, \ldots, x_n \in \overline{\mathbb{R}}$, then

$$\bigvee_{k=1}^{n} x_k = \max\{x_1, \ldots, x_n\} \quad \text{and} \quad \bigwedge_{k=1}^{n} x_k = \min\{x_1, \ldots, x_n\}.$$

Proposition 1.1. *If $(x_n)_{n\in\mathbb{N}}$ is a sequence in $\overline{\mathbb{R}}$, then*

$$\sup_n x_n = \lim_{n\to\infty} \bigvee_{k=1}^{n} x_k \quad \textit{and} \quad \inf_n x_n = \lim_{n\to\infty} \bigwedge_{k=1}^{n} x_k. \tag{1.1}$$

Proof. Only the first limit in (1.1) will be proved, since the proof of the second limit is quite similar.

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\overline{\mathbb{R}}$, and let $L = \lim_{n\to\infty} \bigvee_{k=1}^{n} x_k$. Either (x_n) has an upper bound in $\mathbb{R}$ or it doesn't. If (x_n) doesn't have an upper bound in $\mathbb{R}$, then it is easy to see that $L = \sup_n x_n = +\infty$.

Suppose (x_n) has an upper bound in $\mathbb{R}$, so that $\sup_n x_n \in \mathbb{R}$ by the Completeness Axiom. The limit L exists in $\mathbb{R}$ since the sequence $(\bigvee_{k=1}^{n} x_k)_{n\in\mathbb{N}}$ is monotone increasing and bounded above by the real number $\sup_n x_n$. Clearly $L \leq \sup_n x_n$. Let $\epsilon > 0$. Then there exists some $m \in \mathbb{N}$ such that $\sup_n x_n - \epsilon < x_m \leq \sup_n x_n$, so

$$\sup_n x_n - \epsilon < \bigvee_{k=1}^{n} x_k \leq \sup_n x_n$$

for all $n \geq m$, and hence $\sup_n x_n - \epsilon \leq L \leq \sup_n x_n$. This shows that $L \geq \sup_n x_n$, and therefore $L = \sup_n x_n$ once again. ■

Definition 1.2. *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\overline{\mathbb{R}}$. We define the **limit superior** of (x_n) to be*

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty}\left(\sup_{k\geq n} x_k\right),$$

*and the **limit inferior** to be*

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty}\left(\inf_{k\geq n} x_k\right).$$

Alternative symbols for limit superior $\limsup_n$ *and* $\limsup$*, with similar alternatives for limit inferior.*

Given a sequence of functions $(f_n : X \to \overline{\mathbb{R}})_{n\in\mathbb{N}}$, we define functions $\sup_n f_n : X \to \overline{\mathbb{R}}$ and $\inf_n f_n : X \to \overline{\mathbb{R}}$ by

$$(\sup_n f_n)(x) = \sup_{n\geq 1} f_n(x) \quad \text{and} \quad (\inf_n f_n)(x) = \inf_{n\geq 1} f_n(x),$$

and we define $\limsup_n f_n : X \to \overline{\mathbb{R}}$ and $\liminf_n f_n : X \to \overline{\mathbb{R}}$ by

$$(\limsup_n f_n)(x) = \limsup_{n\to\infty} f_n(x) \quad \text{and} \quad (\liminf_n f_n)(x) = \liminf_{n\to\infty} f_n(x).$$

Let X be a metric space. The **upper limit** and **lower limit** of a function $f : X \to \overline{\mathbb{R}}$ is defined to be

$$\limsup_{x\to a} f(x) = \lim_{\epsilon\to 0}\left(\sup\left\{f(x) : x \in \mathrm{Dom}(f) \cap B'_\epsilon(a)\right\}\right)$$

and

$$\liminf_{x\to a} f(x) = \lim_{\epsilon\to 0}\left(\inf\left\{f(x) : x \in \mathrm{Dom}(f) \cap B'_\epsilon(a)\right\}\right),$$

respectively.

We say a sequence of sets $(A_n)_{n\in\mathbb{N}}$ (also written as $(A_n)_{n=1}^\infty$) **increases** to A, and write $A_n \uparrow A$, if $A_n \subseteq A_{n+1}$ for all $n \in \mathbb{N}$ and

$$A = \bigcup_{n=1}^\infty A_n.$$

We say $(A_n)_{n=1}^\infty$ **decreases** to A, and write $A_n \downarrow A$, if $A_{n+1} \subseteq A_n$ for all $n \in \mathbb{N}$ and

$$A = \bigcap_{n=1}^\infty A_n.$$

We define the **upper limit** and **lower limit** of a sequence of sets $(A_n)_{n\in\mathbb{N}}$ to be

$$\limsup_n A_n = \bigcap_{n=1}^\infty \bigcup_{k=n}^\infty A_k \quad \text{and} \quad \liminf_n A_n = \bigcup_{n=1}^\infty \bigcap_{k=n}^\infty A_k,$$

and if

$$\limsup_n A_n = \liminf_n A_n = A,$$

then we say A is the **limit** of $(A_n)_{n\in\mathbb{N}}$ and write $\lim_n A_n = A$. The following proposition is straightforward to verify.

Proposition 1.3. *Given a sequence of sets* $(A_n)_{n\in\mathbb{N}}$*,*

$$\limsup_n A_n = \{x : x \in A_n \text{ for infinitely many } n\}$$

and

$$\liminf_n A_n = \{x : x \in A_n \text{ for all but finitely many } n\}$$

Given a set X, the **power set** of X, denoted by $\mathcal{P}(X)$, is defined to be the collection of all subsets of X; that is,

$$\mathcal{P}(X) = \{A : A \subseteq X\}.$$

If $A \in \mathcal{P}(X)$, then we define $A^c = X - A$. In the following theorem the symbol $\sqcup$ denotes the union of disjoint sets.

Theorem 1.4. *If* $(A_n)_{n\in\mathbb{N}}$ *is a sequence in* $\mathcal{P}(X)$*, then*

$$\bigcup_{n=1}^{\infty} A_n = \bigsqcup_{n=1}^{\infty} (A_1^c \cap \cdots \cap A_{n-1}^c \cap A_n).$$

If in addition $(A_n)_{n\in\mathbb{N}}$ *is an increasing sequence and* $A_0 = \varnothing$*, then*

$$\bigcup_{n=1}^{\infty} A_n = \bigsqcup_{n=1}^{\infty} (A_n - A_{n-1}).$$

Example 1.5. Show that

$$(\limsup_n A_n) \cup (\limsup_n B_n) = \limsup_n (A_n \cup B_n) \tag{1.2}$$

and

$$(\liminf_n A_n) \cap (\liminf_n B_n) = \liminf_n (A_n \cap B_n) \tag{1.3}$$

Solution. Suppose $x \notin (\limsup A_n) \cup (\limsup B_n)$, so that $x \notin \limsup A_n$ and $x \notin \limsup B_n$. Then there exists some $n_1, n_2 \in \mathbb{N}$ such that

$$x \notin \bigcup_{k=n_1}^{\infty} A_k \quad \text{and} \quad x \notin \bigcup_{k=n_2}^{\infty} B_k.$$

Letting $m = \max\{n_1, n_2\}$, it follows that

$$x \notin \left(\bigcup_{k=m}^{\infty} A_k\right) \cup \left(\bigcup_{k=m}^{\infty} B_k\right) = \bigcup_{k=m}^{\infty} (A_k \cup B_k),$$

and therefore

$$x \notin \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} (A_k \cup B_k) = \limsup (A_n \cup B_n).$$

This shows that

$$\limsup (A_n \cup B_n) \subseteq (\limsup A_n) \cup (\limsup B_n).$$

Now suppose that $x \in (\limsup A_n) \cup (\limsup B_n)$. In particular assume $x \in \limsup A_n$. For any $n \in \mathbb{N}$ we have $x \in \bigcup_{k=n}^{\infty} A_k$, which implies $x \in \bigcup_{k=n}^{\infty}(A_k \cup B_k)$, and hence

$$x \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty}(A_k \cup B_k) = \limsup(A_n \cup B_n).$$

The same conclusion obtains if we assume $x \in \limsup B_n$, and therefore

$$(\limsup A_n) \cup (\limsup B_n) \subseteq \limsup(A_n \cup B_n).$$

This verifies (1.2).

Next, suppose $x \in (\liminf A_n) \cap (\liminf B_n)$, so there exist $n_1, n_2 \in \mathbb{N}$ such that

$$x \in \bigcap_{k=n_1}^{\infty} A_k \quad \text{and} \quad x \in \bigcap_{k=n_2}^{\infty} B_k,$$

and so if $m = \max\{n_1, n_2\}$ we have

$$x \in \left(\bigcap_{k=m}^{\infty} A_k\right) \cap \left(\bigcap_{k=m}^{\infty} B_k\right) = \bigcap_{k=m}^{\infty}(A_k \cap B_k).$$

Hence

$$x \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty}(A_k \cap B_k),$$

and we've shown that

$$(\liminf A_n) \cap (\liminf B_n) \subseteq \liminf(A_n \cap B_n).$$

Now suppose that $x \in \liminf(A_n \cap B_n)$, so there exists some $m \in \mathbb{N}$ such that

$$x \in \bigcap_{k=m}^{\infty}(A_k \cap B_k),$$

which implies that

$$x \in \bigcap_{k=m}^{\infty} A_k \quad \text{and} \quad x \in \bigcap_{k=m}^{\infty} B_k,$$

and hence

$$x \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k = \liminf A_n \quad \text{and} \quad x \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} B_k = \liminf B_n.$$

Thus $x \in (\liminf A_n) \cap (\liminf B_n)$, giving

$$(\liminf A_n) \cap (\liminf B_n) \supseteq \liminf(A_n \cap B_n),$$

which verifies (1.3). ■

1.2 – Algebras and Borel Sets

Definition 1.6. *Let X be a set, and let $\mathcal{A} \subseteq \mathcal{P}(X)$. Then $\mathcal{A}$ is an* ***algebra on*** *X if the following properties hold.*

1. *$X \in \mathcal{A}$.*
2. *If $A \in \mathcal{A}$, then $X \setminus A \in \mathcal{A}$.*
3. *For all $m \in \mathbb{N}$, $A_1, \dots, A_m \in \mathcal{A}$ implies $\bigcup_{n=1}^m A_n \in \mathcal{A}$.*

If, in addition, $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$ for any countable collection $\{A_n : n \in \mathbb{N}\} \subseteq \mathcal{A}$, then $\mathcal{A}$ is a ***σ-algebra on*** *X.*

It can be seen from Definition 1.6 that $X = \bigcup \mathcal{A}$. Thus if a collection of sets $\mathcal{A}$ is simply given to be an algebra, then the associated "universal set" X is understood to be $\bigcup \mathcal{A}$, and so $\mathcal{A}$ is specifically an algebra on $\bigcup \mathcal{A}$. The symbol A^c may also be used to denote $X \setminus A$.

Proposition 1.7.

1. *If $\mathcal{A}$ is an algebra, then $A_1, \dots, A_m \in \mathcal{A}$ implies $\bigcap_{n=1}^m A_n \in \mathcal{A}$ for all $m \in \mathbb{N}$.*
2. *If $\mathcal{S}$ is a σ-algebra, then $\{S_n : n \in \mathbb{N}\} \subseteq \mathcal{S}$ implies $\bigcap_{n=1}^\infty S_n \in \mathcal{S}$.*

Proof. Suppose $\mathcal{S}$ is a σ-algebra and $\{S_n : n \in \mathbb{N}\} \subseteq \mathcal{S}$. Then $\{S_n^c : n \in \mathbb{N}\} \subseteq \mathcal{S}$ as well, and so $\bigcup_{n=1}^\infty S_n^c \in \mathcal{S}$. Now,

$$\bigcup_{n=1}^\infty S_n^c \in \mathcal{S} \;\;\Rightarrow\;\; \left(\bigcap_{n=1}^\infty S_n\right)^c \in \mathcal{S} \;\;\Rightarrow\;\; \bigcap_{n=1}^\infty S_n \in \mathcal{S},$$

which proves part (2). The proof of part (1) is similar. ■

Definition 1.8. *Let X be a set, and let $\mathcal{E} \subseteq \mathcal{P}(X)$. Then $\mathcal{E}$ is an* ***elementary family on*** *X if the following properties hold.*

1. *$\varnothing \in \mathcal{E}$.*
2. *If $E, F \in \mathcal{E}$, then $E \cap F \in \mathcal{E}$.*
3. *If $E \in \mathcal{E}$, then $X \setminus E$ is a finite disjoint union of elements of $\mathcal{E}$.*

Proposition 1.9. *Let $\mathcal{E}$ be an elementary family on X. If $\mathcal{A}$ is the collection of all finite disjoint unions of elements of $\mathcal{E}$, then $\mathcal{A}$ is an algebra on X.*

Proof. Suppose $\mathcal{A}$ is the collection of all finite disjoint unions of elements of $\mathcal{E}$. Since $\varnothing \in \mathcal{E}$ implies $X = \varnothing^c$ is a disjoint union of elements of $\mathcal{E}$, we see that $X \in \mathcal{A}$.

Suppose $E, F \in \mathcal{E}$. For some $m \in \mathbb{N}$ there exist disjoint $G_1, \dots, G_n \in \mathcal{E}$ such that $E^c = \bigsqcup_{k=1}^n G_k$, and thus

$$E \cup F = E \sqcup (F \setminus E) = E \sqcup (F \cap E^c) = E \sqcup \left(\bigsqcup_{k=1}^n (F \cap G_k)\right) \in \mathcal{A}$$

since $F \cap G_k \in \mathcal{E}$ for each $1 \le k \le n$. By induction it follows that $\bigcup_{k=1}^n E_k \in \mathcal{A}$ whenever $E_1, \ldots, E_n \in \mathcal{E}$ for some $n \in \mathbb{N}$. Now let $A_1, \ldots, A_n \in \mathcal{A}$, so for each $1 \le k \le n$ there exist $E_1^k, \ldots, E_{m_k}^k \in \mathcal{E}$ such that $A_k = \bigsqcup_{j=1}^{m_k} E_j^k$. Then $\bigcup_{k=1}^n A_k$ is a finite union of elements of $\mathcal{E}$,

$$\bigcup_{k=1}^n A_k = \bigcup_{k=1}^n \left(\bigsqcup_{j=1}^{m_k} E_j^k \right),$$

and so $\bigcup_{k=1}^n A_k \in \mathcal{A}$.

Finally, if $A \in \mathcal{A}$ is such that $A \in \mathcal{E}$ as well, then A^c is a disjoint union of elements of $\mathcal{E}$ and we have $A^c \in \mathcal{A}$. That is, if $A \in \mathcal{A}$ is a disjoint union of *one* element of $\mathcal{E}$, then $A^c \in \mathcal{A}$. For an arbitrary fixed $n \in \mathbb{N}$, suppose $A^c \in \mathcal{A}$ for any $A \in \mathcal{A}$ that is a disjoint union of n elements of $\mathcal{E}$. Now suppose $A \in \mathcal{A}$ is a disjoint union of $n+1$ elements of $\mathcal{E}$: $A = \bigsqcup_{k=1}^{n+1} E_k$. Let $B = \bigsqcup_{k=1}^n E_k$ and $C = E_{n+1}$. Then $B^c \in \mathcal{A}$ by our inductive hypothesis, and $C^c \in \mathcal{A}$ by the base case. It follows that B^c and C^c are finite disjoint unions of elements of $\mathcal{E}$: $B^c = \bigsqcup_{i=1}^{\ell} F_i$ and $C^c = \bigsqcup_{j=1}^m G_j$. Now,

$$A^c = (B \sqcup C)^c = B^c \cap C^c = \bigsqcup_{i=1}^{\ell} (F_i \cap C^c) = \bigsqcup_{i=1}^{\ell} \bigsqcup_{j=1}^{m} (F_i \cap G_j),$$

where $F_i \cap G_j \in \mathcal{A}$ for all i and j. Since A^c is a finite union of elements of $\mathcal{E}$, we conclude that $A^c \in \mathcal{A}$. Therefore $\mathcal{A}$ is closed under complementation by the principle of induction, and the proof that $\mathcal{A}$ is an algebra on X is done. ■

Example 1.10. The collection $\mathcal{R} \subseteq \mathcal{P}(\mathbb{R}^2)$ of all **rectangles** in $\mathbb{R}^2$,

$$\mathcal{R} = \{I \times J : I, J \subseteq \mathbb{R} \text{ are intervals}\},$$

is an elementary family. So long as we regard singletons $\{a\}$ as also intervals, on purely geometrical grounds it is clear that the intersection of two rectangles $I_1 \times J_1$ and $I_2 \times J_2$ is either $\varnothing$ or another rectangle. Moreover,

$$(I \times J)^c = (I^c \times J) \sqcup (I \times J^c) \sqcup (I^c \times J^c),$$

where I^c and J^c are each either a single interval or a disjoint union of two intervals, and so $(I \times J)^c$ can be expressed as a disjoint union of rectangles. For instance,

$$\big([a,b] \times [c,d]\big)^c = \big((-\infty, a) \times \mathbb{R}\big) \sqcup \big([a,b] \times (-\infty, c)\big) \sqcup \big([a,b] \times (d, \infty)\big) \sqcup \big((b, \infty) \times \mathbb{R}\big)$$

for any $a, b, c, d \in \mathbb{R}$. ■

For a set X let $\mathcal{C} \subseteq \mathcal{P}(X)$. We define the **$\sigma$-algebra on X generated by** $\mathcal{C}$ to be

$$\sigma_X(\mathcal{C}) = \bigcap \big\{ \mathcal{S} \supseteq \mathcal{C} : \mathcal{S} \text{ is a } \sigma\text{-algebra on } X \big\}.$$

Thus $\sigma_X(\mathcal{C})$ is the "smallest" σ-algebra on X containing the collection $\mathcal{C}$. Often $\sigma_X(\mathcal{C})$ is written as $\sigma(\mathcal{C})$ when $\mathcal{C}$ is given to be a subcollection of $\mathcal{P}(X)$, or when the context of a discussion makes clear that X is the universal set.

Definition 1.11. *Let $(X, \mathcal{T})$ be a topological space. The collection of* ***Borel sets in*** *X is*

$$\mathcal{B}(X) = \sigma_X(\mathcal{T}).$$

That is, the Borel sets in X are the sets in the smallest σ-algebra on X containing all the open subsets of X.

Example 1.12 (Borel Sets in $\mathbb{R}$). The σ-algebra of Borel sets of $\mathbb{R}$, denoted by $\mathcal{B}(\mathbb{R})$, is the σ-algebra generated by the open subsets of $\mathbb{R}$. Let

$$\mathcal{O} = \{(a, b) : a, b \in \mathbb{R}\},$$

which is a proper subset of the collection of open subsets of $\mathbb{R}$. Since any open set in $\mathbb{R}$ can be characterized as a countable union of bounded open intervals in $\mathbb{R}$, we find that $\mathcal{O}$ (the collection of all bounded open intervals) also generates the Borel sets of $\mathbb{R}$: $\sigma(\mathcal{O}) = \mathcal{B}(\mathbb{R})$.

A bounded **h-interval** in $\mathbb{R}$ is an interval of the form $(a, b]$, where $a, b \in \mathbb{R}$ with $a < b$. Since

$$(a, b] = \bigcap_{n=1}^{\infty} (a, b + 1/n),$$

we find with the help of Proposition 1.7 that the collection

$$\mathcal{H} = \{(a, b] : a, b \in \mathbb{R}\}$$

of all bounded h-intervals in $\mathbb{R}$ is such that $\sigma(\mathcal{H}) = \mathcal{B}(\mathbb{R})$. Also the collection

$$\mathcal{C} = \{[a, b] : a, b \in \mathbb{R}\}$$

of all bounded closed intervals in $\mathbb{R}$ generate the Borel sets of $\mathbb{R}$. Thus we have

$$\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{O}) = \sigma(\mathcal{H}) = \sigma(\mathcal{C}),$$

and there are other generating sets for $\mathcal{B}(\mathbb{R})$ besides. ■

Example 1.13 (Borel Sets in $\overline{\mathbb{R}}$). By definition, $\mathcal{B}(\overline{\mathbb{R}}) = \sigma(\mathcal{T}_{\overline{\mathbb{R}}})$, where $\mathcal{T}_{\overline{\mathbb{R}}}$ is the collection of open subsets of $\overline{\mathbb{R}}$. The question naturally arises: what are the open subsets of $\overline{\mathbb{R}}$? The customary choice for a topology on $\overline{\mathbb{R}}$ may be most easily described as the topology generated by the basis

$$\mathcal{U} = \big\{(a, b) : a, b \in \mathbb{R}\big\} \cup \big\{[-\infty, b) : b \in \mathbb{R}\big\} \cup \big\{(a, \infty] : a \in \mathbb{R}\big\}.$$

That is, the basis for the topology $\mathcal{T}_{\overline{\mathbb{R}}}$ on $\overline{\mathbb{R}}$ consists of the basis for the usual topology $\mathcal{T}$ on $\mathbb{R}$ together with all intervals of the form $[-\infty, b)$ and $(a, \infty]$ for $a, b \in \mathbb{R}$. From this it can be shown that $\mathcal{B}(\overline{\mathbb{R}}) = \sigma(\overline{\mathcal{H}})$, where

$$\overline{\mathcal{H}} = \{(a, b] : a, b \in \overline{\mathbb{R}}\}.$$

Now consider the collection

$$\overline{\mathcal{O}} = \{(a, \infty] : a \in \mathbb{R}\}.$$

Does $\sigma(\overline{\mathcal{O}}) = \mathcal{B}(\overline{\mathbb{R}})$? For any $-\infty < a < b < \infty$ we have $(a, \infty], (b, \infty] \in \sigma(\mathcal{O})$, and thus

$$(b, \infty] \in \sigma(\overline{\mathcal{O}}) \;\;\Rightarrow\;\; [-\infty, b] = (b, \infty]^c \in \sigma(\overline{\mathcal{O}})$$

$$\Rightarrow \quad (a,b] = [-\infty,b] \cap (a,\infty] \in \sigma(\overline{\mathcal{O}}).$$

In addition, for each $b \in \overline{\mathbb{R}}$ we have $(-n,b] \in \sigma(\overline{\mathcal{O}})$ for all $n \in \mathbb{N}$, and hence

$$(-\infty,b] = \bigcup_{n=1}^{\infty}(-n,b] \in \sigma(\overline{\mathcal{O}}).$$

This shows that $\sigma(\overline{\mathcal{O}}) \supseteq \overline{\mathcal{H}}$, and hence $\sigma(\overline{\mathcal{O}}) \supseteq \sigma(\overline{\mathcal{H}}) = \mathcal{B}(\overline{\mathbb{R}})$ since $\sigma(\overline{\mathcal{H}})$ is by definition the smallest σ-algebra containing $\overline{\mathcal{H}}$. On the other hand all the elements of $\overline{\mathcal{O}}$ are open sets in $\overline{\mathbb{R}}$, implying that $\overline{\mathcal{O}} \subseteq \mathcal{B}(\overline{\mathbb{R}})$ and hence $\sigma(\overline{\mathcal{O}}) \subseteq \mathcal{B}(\overline{\mathbb{R}})$. Therefore $\sigma(\overline{\mathcal{O}}) = \mathcal{B}(\overline{\mathbb{R}})$.

Since the complement (relative to $\overline{\mathbb{R}}$) of each element of $\overline{\mathcal{O}}$ is of the form $[-\infty,a]$, where $a \in \mathbb{R}$, it is immediate that the collection $\{[-\infty,a] : a \in \mathbb{R}\}$ also generates $\mathcal{B}(\overline{\mathbb{R}})$, which straightaway implies that the larger collection

$$\overline{\mathcal{C}} = \{[a,b] : a,b \in \overline{\mathbb{R}}\}$$

generates $\mathcal{B}(\overline{\mathbb{R}})$. ∎

If $\mathcal{C} \subseteq \mathcal{P}(X)$ and $A \subseteq X$, then we define

$$\mathcal{C} \cap A = \{B \cap A : B \in \mathcal{C}\}.$$

Thus $\mathcal{C} \cap A \subseteq \mathcal{P}(A)$, and so we may consider the σ-algebra on A generated by $\mathcal{C} \cap A$.

Proposition 1.14. *If $\mathcal{S}$ is a σ-algebra on X and $A \subseteq X$, then $\mathcal{S} \cap A$ is a σ-algebra on A.*

Proof. Suppose $\mathcal{S}$ is a σ-algebra on X and $A \subseteq X$. Clearly $A \in \mathcal{S} \cap A$ since $X \in \mathcal{S}$. Let $E \in \mathcal{S} \cap A$, so $E = S \cap A$ for some $S \in \mathcal{S}$. Now, $X \setminus S \in \mathcal{S}$, and so

$$A \setminus E = A \cap (X \setminus E) = A \cap \big[(X \setminus S) \cup (X \setminus A)\big] = (X \setminus S) \cap A \in \mathcal{S} \cap A.$$

Finally, let $\{E_n : n \in \mathbb{N}\} \subseteq \mathcal{S} \cap A$, so for each n there is some $S_n \in \mathcal{S}$ such that $E_n = S_n \cap A$. Since $\bigcup_{n=1}^{\infty} S_n \in \mathcal{S}$, it follows that

$$\bigcup_{n=1}^{\infty} E_n = \left(\bigcup_{n=1}^{\infty} S_n\right) \cap A \in \mathcal{S} \cap A,$$

and we're done. ∎

Proposition 1.15. *If $\mathcal{C} \subseteq \mathcal{P}(X)$ and $A \subseteq X$, then*

$$\sigma_A(\mathcal{C} \cap A) = \sigma_X(\mathcal{C}) \cap A.$$

Proof. It is clear that $\mathcal{C} \cap A \subseteq \sigma_X(\mathcal{C}) \cap A$, and since $\sigma_X(\mathcal{C}) \cap A$ is a σ-algebra on A containing $\mathcal{C} \cap A$, it follows that $\sigma_A(\mathcal{C} \cap A) \subseteq \sigma_X(\mathcal{C}) \cap A$.

Define $\mathcal{S} \subseteq \sigma_X(\mathcal{C})$ by

$$\mathcal{S} = \{B \in \sigma_X(\mathcal{C}) : B \cap A \in \sigma_A(\mathcal{C} \cap A)\}.$$

We wish to show that $\mathcal{C} \subseteq \mathcal{S}$ and $\mathcal{S}$ is a σ-algebra.

Let $B \in \mathcal{C}$. Then $B \in \sigma_X(\mathcal{C})$ and also

$$B \cap A \in \mathcal{C} \cap A \subseteq \sigma_A(\mathcal{C} \cap A),$$

which shows that $B \in \mathcal{S}$ and hence $\mathcal{C} \subseteq \mathcal{S}$.

Let $(B_n)_{n\in\mathbb{N}}$ be a sequence of sets in $\mathcal{S}$, so $B_n \in \sigma_X(\mathcal{C})$ and $B_n \cap A \in \sigma_A(\mathcal{C} \cap A)$ for all n. Since

$$\bigcup_{n=1}^{\infty} B_n \in \sigma_X(\mathcal{C})$$

and

$$\left(\bigcup_{n=1}^{\infty} B_n\right) \cap A = \bigcup_{n=1}^{\infty} (B_n \cap A) \in \sigma_A(\mathcal{C} \cap A),$$

it follows that $\bigcup_{n=1}^{\infty} B_n \in \mathcal{S}$.

If $B \in \mathcal{S}$, so that $B \in \sigma_X(\mathcal{C})$ and $B \cap A \in \sigma_A(\mathcal{C} \cap A)$, then we have $B^c \in \sigma_X(\mathcal{C})$ and $(B \cap A)^c \in \sigma_A(\mathcal{C} \cap A)$ also, with the latter implying that

$$B^c \cap A = (B^c \cup A^c) \cap A = (B \cap A)^c \cap A \in \sigma_A(\mathcal{C} \cap A),$$

and hence $B^c \in \mathcal{S}$.

Finally, $X \in \sigma_X(\mathcal{C})$ and $X \cap A = A \in \sigma_A(\mathcal{C} \cap A)$ shows that $X \in \mathcal{S}$, and $\mathcal{S}$ is now verified to be a σ-algebra.

Since $\mathcal{C} \subseteq \mathcal{S}$ and $\mathcal{S}$ is a σ-algebra, we have $\sigma_X(\mathcal{C}) \subseteq \mathcal{S}$, and hence $\sigma_X(\mathcal{C}) = \mathcal{S}$. So $B \cap A \in \sigma_A(\mathcal{C} \cap A)$ for all $B \in \sigma_X(\mathcal{C})$, and therefore $\sigma_X(\mathcal{C}) \cap A \subseteq \sigma_A(\mathcal{C} \cap A)$. ■

Example 1.16. Given a topological space $(X, \mathcal{T})$ and a set $A \subseteq X$, the **subspace topology on** A (relative to X) is the topological space $(A, \mathcal{T}_A)$, where we define

$$\mathcal{T}_A = \mathcal{T} \cap A = \{U \cap A : U \in \mathcal{T}\}.$$

It then follows that the Borel sets in A is the collection

$$\mathcal{B}(A) = \sigma_A(\mathcal{T} \cap A) = \sigma_X(\mathcal{T}) \cap A = \mathcal{B}(X) \cap A. \tag{1.4}$$

by Proposition 1.15. ■

Remark. Let (X, d) be a metric space. For each $x_0 \in X$ and $r > 0$ define the **open ball** with **center** x_0 and **radius** r to be the set

$$B_r(x_0) = \{x \in X : d(x_0, x) < r\}.$$

As discussed in elementary analysis, the metric d induces a topology $\mathcal{T}$ on X as follows: $U \in \mathcal{T}$ iff for each $u \in U$ there exists some $r > 0$ such that $B_r(u) \subseteq U$. Let $A \subseteq X$. A fact from point-set topology states that the topology on the metric space (A, d) is precisely the same as $\mathcal{T} \cap A$; that is, the topology the metric d induces on A equals the subspace topology on A relative to X.

Example 1.17. Recall the collection $\mathcal{U}$ in Example 1.13 that is defined to be the basis for the topological space $(\overline{\mathbb{R}}, \mathcal{T}_{\overline{\mathbb{R}}})$. Is the subspace topology on $\mathbb{R}$ relative to $\overline{\mathbb{R}}$ the same as the standard topology $\mathcal{T}_{\mathbb{R}}$ on $\mathbb{R}$? That is, do we have $\mathcal{T}_{\mathbb{R}} = \mathcal{T}_{\overline{\mathbb{R}}} \cap \mathbb{R}$?

In general, if $\mathcal{B}$ is a basis for a topological space $(X, \mathcal{T})$, and $A \subseteq X$, then $\mathcal{B} \cap A$ is a basis for the topological subspace $(A, \mathcal{T} \cap A)$. Thus $\mathcal{U} \cap \mathbb{R}$ is a basis for $(\mathbb{R}, \mathcal{T}_{\overline{\mathbb{R}}} \cap \mathbb{R})$, the subspace topology on $\mathbb{R}$ relative to $\overline{\mathbb{R}}$. It is clear that $\mathcal{U} \cap \mathbb{R}$ contains the collection $\{(a, b) : a, b \in \mathbb{R}\}$, which is a basis for $\mathcal{T}_{\mathbb{R}}$, and so $\mathcal{T}_{\mathbb{R}} \subseteq \mathcal{T}_{\overline{\mathbb{R}}} \cap \mathbb{R}$. On the other hand, any $U \in \mathcal{T}_{\overline{\mathbb{R}}}$ is a union of elements of $\mathcal{U}$, whereupon it follows that $U \cap \mathbb{R}$ is a union of open intervals in $\mathbb{R}$ and thus $U \cap \mathbb{R} \in \mathcal{T}_{\mathbb{R}}$. This shows that $\mathcal{T}_{\overline{\mathbb{R}}} \cap \mathbb{R} \subseteq \mathcal{T}_{\mathbb{R}}$, and therefore $\mathcal{T}_{\mathbb{R}} = \mathcal{T}_{\overline{\mathbb{R}}} \cap \mathbb{R}$.

Knowing that $\mathcal{T}_{\mathbb{R}} = \mathcal{T}_{\overline{\mathbb{R}}} \cap \mathbb{R}$, we can now employ (1.4) in the previous example to conclude that

$$\mathcal{B}(\mathbb{R}) = \mathcal{B}(\overline{\mathbb{R}}) \cap \mathbb{R},$$

which will be useful later on. ■

Problem 1.18 (PMT 1.2.6). Let $f : X \to Y$, and let $\mathcal{C} \subseteq \mathcal{P}(Y)$. Show that

$$\sigma(f^{-1}(\mathcal{C})) = f^{-1}(\sigma(\mathcal{C})),$$

where we define

$$f^{-1}(\mathcal{C}) = \{f^{-1}(A) : A \in \mathcal{C}\} \quad \text{and} \quad f^{-1}\big(\sigma(\mathcal{C})\big) = \{f^{-1}(A) : A \in \sigma(\mathcal{C})\}.$$

Solution. If $B \in f^{-1}(\mathcal{C})$, then $B = f^{-1}(A)$ for some $A \in \mathcal{C} \subseteq \sigma(\mathcal{C})$, which implies that $B \in f^{-1}(\sigma(\mathcal{C}))$. Thus $f^{-1}(\mathcal{C}) \subseteq f^{-1}(\sigma(\mathcal{C}))$, and since $f^{-1}(\sigma(\mathcal{C}))$ is a σ-algebra, we conclude that $\sigma(f^{-1}(\mathcal{C})) \subseteq f^{-1}(\sigma(\mathcal{C}))$.

Define

$$\mathcal{S} = \{B \in \sigma(\mathcal{C}) : f^{-1}(B) \in \sigma(f^{-1}(\mathcal{C}))\},$$

where of course $\mathcal{S} \subseteq \sigma(\mathcal{C})$. We wish to show that $\mathcal{C} \subseteq \mathcal{S}$ and $\mathcal{S}$ is a σ-algebra, from which it will follow that $\sigma(\mathcal{C}) \subseteq \mathcal{S}$, and therefore $\mathcal{S} = \sigma(\mathcal{C})$. This in turn will imply that $f^{-1}(\sigma(\mathcal{C})) \subseteq \sigma(f^{-1}(\mathcal{C}))$, and the proof will be done.

Let $B \in \mathcal{C}$. Then $B \in \sigma(\mathcal{C})$, and also $f^{-1}(B) \in f^{-1}(\mathcal{C}) \subseteq \sigma(f^{-1}(\mathcal{C}))$, so that $B \in \mathcal{S}$ and we have $\mathcal{C} \subseteq \mathcal{S}$.

Let $(B_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{S}$, so $B_n \in \sigma(\mathcal{C})$ and $f^{-1}(B_n) \in \sigma(f^{-1}(\mathcal{C}))$ for all n. Since

$$\bigcup_{n=1}^{\infty} B_n \in \sigma(\mathcal{C})$$

and

$$f^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right) = \bigcup_{n=1}^{\infty} f^{-1}(B_n) \in \sigma(f^{-1}(\mathcal{C})),$$

it follows that $\bigcup_{n=1}^{\infty} B_n \in \mathcal{S}$ and $\mathcal{S}$ is countably additive.

If $B \in \mathcal{S}$, so that $B \in \sigma(\mathcal{C})$ and $f^{-1}(B) \in \sigma(f^{-1}(\mathcal{C}))$, then we have $B^c \in \sigma(\mathcal{C})$ and $f^{-1}(B^c) = [f^{-1}(B)]^c \in \sigma(f^{-1}(\mathcal{C}))$ also, showing that $B^c \in \mathcal{S}$ and $\mathcal{S}$ is closed under complementation.

Finally, $Y \in \sigma(\mathcal{C})$, and $f^{-1}(Y) = X \in \sigma(f^{-1}(\mathcal{C}))$ since $f^{-1}(\mathcal{C})$ is a collection of subsets of X. The collection $\mathcal{S}$ is now verified to be a σ-algebra. ■

Problem 1.19 (PMT 1.2.7). Suppose $A \in \mathcal{B}(\mathbb{R})$. Let $\mathcal{T}_{\mathbb{R}}$ be the standard topology on $\mathbb{R}$, so that $\mathcal{T}_{\mathbb{R}} \cap A$ is the usual subspace topology on A. Show that the smallest σ-algebra on A containing the sets open in A is the collection $\mathcal{C} = \{B \in \mathcal{B}(\mathbb{R}) : B \subseteq A\}$. Thus

$$\mathcal{B}(A) = \sigma_A(\mathcal{T}_{\mathbb{R}} \cap A) = \{B \in \mathcal{B}(\mathbb{R}) : B \subseteq A\}.$$

Solution. If $S \in \mathcal{C}$, then $S \in \mathcal{B}(\mathbb{R})$ and $S = S \cap A$, and so $S \in \mathcal{B}(\mathbb{R}) \cap A$. Conversely, if $S \in \mathcal{B}(\mathbb{R}) \cap A$, so that $S = B \cap A$ for some $B \in \mathcal{B}(\mathbb{R})$, then $S \subseteq A$, and also $S \in \mathcal{B}(\mathbb{R})$ since $A \in \mathcal{B}(\mathbb{R})$, and we conclude that $S \in \mathcal{C}$. Therefore

$$\{B \in \mathcal{B}(\mathbb{R}) : B \subseteq A\} = \mathcal{B}(\mathbb{R}) \cap A.$$

Now, by definition $\mathcal{B}(A)$ is $\sigma_A(\mathcal{T}_{\mathbb{R}} \cap A)$, the smallest σ-algebra on A containing the sets open in A, and so by Proposition 1.15 we have

$$\mathcal{B}(A) = \sigma_A(\mathcal{T}_{\mathbb{R}} \cap A) = \sigma_{\mathbb{R}}(\mathcal{T}) \cap A = \mathcal{B}(\mathbb{R}) \cap A,$$

and hence $\mathcal{B}(A) = \mathcal{C}$ as desired. ■

1.3 – Set Functions and Measures

A **set function** is any function $\mu : \mathcal{C} \to \overline{\mathbb{R}}$, where $\mathcal{C} \subseteq \mathcal{P}(X)$. Usually $\mathcal{C}$ will be an algebra or σ-algebra. We say μ is **finitely additive** if

$$\mu\left(\bigsqcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_n)$$

for any finite collection $\{A_1, \dots, A_n\} \subseteq \mathcal{C}$ of disjoint sets such that $\bigsqcup_{k=1}^{n} A_k \in \mathcal{C}$, and **countably additive** if

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

for any countable subcollection $\{A_n : n \in \mathbb{N}\} \subseteq \mathcal{C}$ of disjoint sets such that $\bigsqcup_{n=1}^{\infty} A_n \in \mathcal{C}$. We say μ is **finite** if $\mu(A) \in \mathbb{R}$ for all $A \in \mathcal{C}$, and **bounded** if there exists some $M \in \mathbb{R}$ such that $|\mu(A)| \leq M$ for all $A \in \mathcal{C}$. Finally, if μ is nonnegative and finitely additive, and there exists a countable subcollection $\{A_n : n \in \mathbb{N}\} \subseteq \mathcal{C}$ such that $X = \bigcup_{n=1}^{\infty} A_n$ and $\mu(A_n) \in \mathbb{R}$ for each $n \in \mathbb{N}$, then μ is said to be σ**-finite**.

Theorem 1.20. *Let μ be a finitely additive set function on the algebra $\mathcal{A}$, and let $A, B \in \mathcal{A}$.*

1. $\mu(\varnothing) = 0$.
2. $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$.
3. *If $B \subseteq A$, then*
$$\mu(A) = \mu(B) + \mu(A - B),$$
4. *If $B \subseteq A$ and $\mu(B) \in \mathbb{R}$, then*
$$\mu(A - B) = \mu(A) - \mu(B).$$
5. *If $B \subseteq A$ and $\mu(A - B) \geq 0$, then*
$$\mu(B) \leq \mu(A).$$
6. *If μ is nonnegative, then*
$$\mu\left(\bigcup_{n=1}^{m} A_n\right) \leq \sum_{n=1}^{m} \mu(A_n)$$
for all $m \in \mathbb{N}$ and $A_1, \dots, A_m \in \mathcal{A}$.

Theorem 1.21. *Let μ be a countably additive set function on the σ-algebra $\mathcal{S}$, and let $(A_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{S}$.*

1. *If $A_n \uparrow A$, then $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.*
2. *If $A_n \downarrow A$ and $\mu(A_1) \in \mathbb{R}$, then $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.*

Proof.
Proof of (1). Suppose that $A_n \uparrow A$. The case when $\mu(A_n) = \pm\infty$ for some $n \in \mathbb{N}$ is handled fine in [PMT],[1] so assume $\mu(A_n) \in \mathbb{R}$ for all n. Recalling Theorems 1.4 and 1.20(4), countable

[1] *Probability and Measure Theory,* by Ash and Doléans-Dade, 2nd Edition.

additivity gives

$$\mu(A) = \mu\left(\bigsqcup_{n=1}^{\infty}(A_n - A_{n-1})\right) = \sum_{n=1}^{\infty}\mu(A_n - A_{n-1})$$
$$= \sum_{n=1}^{\infty}\big[\mu(A_n) - \mu(A_{n-1})\big] = \lim_{n\to\infty}\sum_{k=1}^{n}\mu(A_k - A_{k-1}) = \lim_{n\to\infty}\mu(A_n),$$

as desired. ■

Definition 1.22. *Let μ be a set function on an algebra $\mathcal{A}$, and let $A \in \mathcal{A}$. We say μ is* ***continuous from below at*** *A if $\lim_{n\to\infty}\mu(A_n) = \mu(A)$ for every sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ such that $A_n \uparrow A$. We say μ is* ***continuous from above at*** *A if $\lim_{n\to\infty}\mu(A_n) = \mu(A)$ for every sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ such that $A_n \downarrow A$.*

We say μ is ***continuous from below on*** *$\mathcal{A}$ if μ is continuous from below at every $A \in \mathcal{A}$, and we say μ is* ***continuous from above on*** *$\mathcal{A}$ if μ is continuous from above at every $A \in \mathcal{A}$.*

Theorem 1.23. *Suppose μ is a finitely additive set function on an algebra $\mathcal{A}$. If μ is continuous from below on $\mathcal{A}$, then μ is countably additive on $\mathcal{A}$. If μ is continuous from above at $\varnothing$, then μ is countably additive on $\mathcal{A}$.*

Problem 1.24 (PMT 1.2.4). Let $\mathcal{A}$ be the algebra of finite disjoint unions of right-semiclosed intervals of $\mathbb{R}$, and define the set function $\mu : \mathcal{A} \to \overline{\mathbb{R}}$ as follows for $a, b \in \mathbb{R}$:

$$\mu(-\infty, b] = b, \quad \mu(a, b] = b - a, \quad \mu(a, \infty) = -a, \quad \mu(\mathbb{R}) = 0,$$

and

$$\mu\left(\bigsqcup_{n=1}^{m} I_n\right) = \sum_{n=1}^{m}\mu(I_n) \tag{1.5}$$

if $I_1, \ldots, I_m$ are disjoint right-semiclosed intervals.

(a) Show that μ is finitely additive but not countably additive on $\mathcal{A}$.

(b) Show that μ is finite but unbounded on $\mathcal{A}$.

Solution.

(a) Let $J_1, \ldots, J_n \in \mathcal{A}$ such that $J_k \cap J_\ell = \varnothing$ whenever $k \neq \ell$. Each J_k is a finite disjoint union of right-semiclosed intervals,

$$J_k = \bigsqcup_{\ell=1}^{m_k} I_{k,\ell},$$

and so

$$\mu(J_k) = \sum_{\ell=1}^{m_k}\mu(I_{k,\ell})$$

by (1.5). Now, again by (1.5),

$$\mu\left(\bigsqcup_{k=1}^{n} J_k\right) = \mu\left(\bigsqcup_{k=1}^{n}\bigsqcup_{\ell=1}^{m_k} I_{k,\ell}\right) = \sum_{k=1}^{n}\sum_{\ell=1}^{m_k}\mu(I_{k,\ell}) = \sum_{k=1}^{n}\mu(J_k),$$

and we conclude that μ is finitely additive.

Next, define the sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ by $A_1 = (-\infty, 1]$, and $A_n = (n-1, n]$ for $n \geq 2$. Then $\bigsqcup_{n=1}^{\infty} A_n = \mathbb{R}$, and so

$$\mu\left(\bigsqcup_{n=1}^{\infty} A_n\right) = \mu(\mathbb{R}) = 0.$$

On the other hand,

$$\begin{aligned}\sum_{n=1}^{\infty} \mu(A_n) &= \lim_{n\to\infty} \sum_{k=1}^{n} \mu(A_k) = \lim_{n\to\infty} \Big(\mu(-\infty, 1] + \mu(1, 2] + \cdots + \mu(n-1, n]\Big) \\ &= \lim_{n\to\infty} \Big[1 + (2-1) + (3-2) + \cdots + \big(n - (n-1)\big)\Big] = \lim_{n\to\infty} n = +\infty.\end{aligned}$$

Thus we see that

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

and so μ is not countably additive.

(b) It is clear that $\mu(A) \in \mathbb{R}$ for any $A \in \mathcal{A}$. However, $\mu(-\infty, n] = n$ for each $n \in \mathbb{N}$ shows that μ is unbounded. ■

Problem 1.25 (PMT 1.2.5). Let μ be a nonnegative, finitely additive set function on the algebra $\mathcal{A}$. If $(A_n)_{n\in\mathbb{N}}$ is a sequence of disjoint sets in $\mathcal{A}$ such that $\bigsqcup_{n=1}^{\infty} A_n \in \mathcal{A}$, show that

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

Solution. By finite additivity and Theorem 1.20(5),

$$\sum_{k=1}^{n} \mu(A_k) = \mu\left(\bigsqcup_{k=1}^{n} A_k\right) \leq \mu\left(\bigsqcup_{n=1}^{\infty} A_n\right)$$

for all $n \in \mathbb{N}$. Therefore

$$\sum_{n=1}^{\infty} \mu(A_n) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(A_k) \leq \mu\left(\bigsqcup_{n=1}^{\infty} A_n\right),$$

where the limit exists in $[0, \infty]$ since the sequence

$$\left(\sum_{k=1}^{n} \mu(A_k)\right)_{n\in\mathbb{N}}$$

is monotone increasing in $[0, \infty]$. ■

Problem 1.26 (PMT 1.2.10). Let μ be a finite measure on the σ-algebra $\mathcal{S}$. If $(A_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{S}$ and $\lim_{n\to\infty} A_n = A$, show that

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

Solution. Suppose $(A_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{S}$ and $\lim_{n\to\infty} A_n = A$. Thus

$$A = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k.$$

Let $B_n = \bigcap_{k=n}^{\infty} A_k$, so $(B_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{S}$ such that $B_n \subseteq B_{n+1}$ and $\bigcup_{n=1}^{\infty} B_n = A$, and so $B_n \uparrow A$. Theorem 1.21(1) implies that $\lim_{n\to\infty}\mu(B_n) = \mu(A)$. Indeed, observing that $B_n \subseteq A_n$ for all $n \in \mathbb{N}$, so that $\mu(B_n) \le \mu(A_n)$ by Theorem 1.20(5), it follows that

$$\mu(A) = \lim_{n\to\infty}\mu(B_n) \le \lim_{n\to\infty}\mu(A_n).$$

Now let $C_n = \bigcup_{k=n}^{\infty} A_k$, so $(C_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{S}$ such that $C_n \supseteq C_{n+1}$ and $\bigcap_{n=1}^{\infty} C_n = A$, and so $C_n \downarrow A$. Thus $\lim_{n\to\infty}\mu(C_n) = \mu(A)$ by Theorem 1.21(2), and since $C_n \supseteq A_n$ for all $n \in \mathbb{N}$, we obtain

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

by Theorem 1.20(5). Therefore $\lim_{n\to\infty}\mu(A_n) = \mu(A)$. ■

Definition 1.27. *A* ***measurable space*** *is a pair $(X, \mathcal{M})$, where X is a set and $\mathcal{M} \subseteq \mathcal{P}(X)$ is a σ-algebra on X whose elements are called* ***measurable sets***. *A set function $\mu : \mathcal{M} \to \overline{\mathbb{R}}$ is a* ***measure on*** *$\mathcal{M}$ if the following properties hold.*

1. *$\mu(A) \ge 0$ for all $A \in \mathcal{M}$.*
2. *$\mu(\varnothing) = 0$.*
3. *Countable Additivity: For every disjoint sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{M}$,*

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

A ***measure space*** *is a triple $(X, \mathcal{M}, \mu)$; that is, a measurable space $(X, \mathcal{M})$ together with a measure μ on $\mathcal{M}$.*

A measure space $(X, \mathcal{M}, \mu)$ is **finite**, **bounded**, or **σ-finite** if $\mu : \mathcal{M} \to \overline{\mathbb{R}}$ is finite, bounded, or σ-finite. If $(X, \mathcal{T})$ is a topological space, then a **Borel measure** is a measure $\mu : \mathcal{B}(X) \to \overline{\mathbb{R}}$, and $(X, \mathcal{B}(X), \mu)$ is a **Borel measure space**. We say a measure space $(X, \mathcal{M}, \mu)$ is a **probability measure space** if $\mu(X) = 1$.

Problem 1.28 (PMT 1.2.12). Show that if μ is a finite measure on a σ-algebra $\mathcal{S}$, there cannot be uncountably many disjoint sets $A \in \mathcal{S}$ such that $\mu(A) > 0$.

Solution. Suppose there are uncountably many disjoint $A \in \mathcal{S}$ such that $\mu(A) > 0$. Thus, if $\mathcal{C}$ is the collection of all such sets in $\mathcal{S}$, then we may write

$$\mathcal{C} = \bigsqcup_{n=0}^{\infty}\left\{A \in \mathcal{S} : \mu(A) \in \left(\tfrac{1}{n+1}, \tfrac{1}{n}\right]\right\},$$

where here we take $\left(\frac{1}{n+1}, \frac{1}{n}\right]$ to be $(1, \infty)$ when $n = 0$. Since $\mathcal{C}$ is uncountable, at least one of the collections

$$\mathcal{C}_m := \left\{A \in \mathcal{S} : \mu(A) \in \left(\tfrac{1}{m+1}, \tfrac{1}{m}\right]\right\}$$

must be at least countably infinite (indeed at least one of the sets must be uncountable). Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of disjoint sets in $\mathcal{C}_m$, so

$$\mu(A_n) \geq \frac{1}{m+1}$$

for all m, and $A = \bigsqcup_{n=1}^{\infty} A_n \in \mathcal{S}$. By the countable additivity of μ we have

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A_n) \geq \sum_{n=1}^{\infty} \frac{1}{m+1} = +\infty,$$

and therefore μ is not a finite measure. ■

1.4 – Extension of Premeasures

The following definition slightly generalizes the notion of a measure $\mu : \mathcal{A} \to \overline{\mathbb{R}}$ to the case when $\mathcal{A}$ is an algebra, but not necessarily a σ-algebra. The substantive difference lies in the statement of the third axiom.

Definition 1.29. *If $\mathcal{A}$ is an algebra, a set function $\mu_0 : \mathcal{A} \to \overline{\mathbb{R}}$ is a **premeasure** on $\mathcal{A}$ if the following properties hold.*

1. $\mu_0(A) \geq 0$ *for all* $A \in \mathcal{A}$.
2. $\mu_0(\varnothing) = 0$.
3. *Countable Additivity: For every disjoint sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ for which $\bigsqcup_{n=1}^\infty A_n \in \mathcal{A}$,*

$$\mu_0\left(\bigsqcup_{n=1}^\infty A_n\right) = \sum_{n=1}^\infty \mu_0(A_n).$$

*For a set X, the triple $(X, \mathcal{A}, \mu_0)$ is a **premeasure space**.*

Theorem 1.30. *If $(X, \mathcal{A}, \mu_0)$ is a premeasure space, then*

$$\mu_0\left(\bigcup_{n=1}^\infty A_n\right) \leq \sum_{n=1}^\infty \mu_0(A_n)$$

for any sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ such that $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$.

The proof of the following proposition is much the same as the proof of Theorem 1.21. Recall that we never allow both $-\infty$ and $+\infty$ to be in the range of a set function $\mu : \mathcal{C} \to \overline{\mathbb{R}}$.

Proposition 1.31. *Let μ be a countably additive set function on the algebra $\mathcal{A}$, and let $(A_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{A}$.*

1. *If $A_n \uparrow A$ and $A \in \mathcal{A}$, then $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.*
2. *If $A_n \uparrow A$ and μ is nonnegative, then $\lim_{n\to\infty} \mu(A_n)$ exists in $\overline{\mathbb{R}}$.*
3. *If $A_n \downarrow A$, $\mu(A_1) \in \mathbb{R}$, and $A \in \mathcal{A}$, then $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.*
4. *If $A_n \downarrow A$, $\mu(A_1) \in \mathbb{R}$, and μ is nonnegative, then $\lim_{n\to\infty} \mu(A_n)$ exists in $\overline{\mathbb{R}}$.*

If a premeasure μ_0 on an algebra $\mathcal{A} \subseteq \mathcal{P}(X)$ is given to be finite, which is to say $\mu_0(A) \in \mathbb{R}$ for all $A \in \mathcal{A}$, then μ_0 is bounded on $\mathcal{A}$ as well. In fact, since for any $A \in \mathcal{A}$ we have $A \subseteq X$, it follows that

$$0 \leq \mu_0(A) \leq \mu_0(X) < \infty$$

by the nonnegativity of μ_0 and Theorem 1.20(5).

Lemma 1.32. *Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ be an algebra, let μ_0 be a finite premeasure on $\mathcal{A}$, and let $(A_n)_{n\in\mathbb{N}}$ and $(A_n')_{n\in\mathbb{N}}$ be sequences in $\mathcal{A}$ such that $A_n \uparrow A$ and $A_n' \uparrow A'$. If $A \subseteq A'$, then*

$$\lim_{n\to\infty} \mu_0(A_n) \leq \lim_{n\to\infty} \mu_0(A_n').$$

Thus if $A' = A$, then

$$\lim_{n\to\infty} \mu_0(A_n) = \lim_{n\to\infty} \mu_0(A_n').$$

Remark. There is no question that the limits featured in the lemma exist as real numbers, since the sequences $(\mu_0(A_n))_{n\in\mathbb{N}}$ and $(\mu_0(A'_n))_{n\in\mathbb{N}}$ are both monotone increasing and bounded above by $\mu_0(\Omega) \in \mathbb{R}$. More explicitly, the sequence $(\mu_0(A_n))_{n\in\mathbb{N}}$ in particular converges to the least upper bound of the set $\{\mu_0(A_n) : n \in \mathbb{N}\} \subseteq \mathbb{R}$, guaranteed to exist in $\mathbb{R}$ by the Completeness Axiom.

Lemma 1.33. *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra, let $\mu_0 : \mathcal{A} \to [0,\infty)$ be a finite premeasure with $\mu_0(X) = r$, and let*

$$\mathcal{G} = \big\{A \in \mathcal{P}(X) : \exists (A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}\,(A_n \uparrow A)\big\}.$$

Define $\mu : \mathcal{G} \to [0,\infty)$ by

$$\mu(A) = \lim_{n\to\infty} \mu_0(A_n),$$

where $(A_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{A}$ such that $A_n \uparrow A$. Then the following hold.

1. *$\mathcal{A} \subseteq \mathcal{G}$, and $\mu|_{\mathcal{A}} = \mu_0$.*
2. *If $A, B \in \mathcal{G}$, then $A \cup B, A \cap B \in \mathcal{G}$, and*

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B).$$

3. *If $A, B \in \mathcal{G}$ with $B \subseteq A$, then $\mu(B) \le \mu(A)$.*
4. *If $(G_n)_{n\in\mathbb{N}} \subseteq \mathcal{G}$ with $G_n \uparrow G$, then $G \in \mathcal{G}$ and*

$$\lim_{n\to\infty} \mu(G_n) = \mu(G).$$

Proof. Note that μ is well-defined on $\mathcal{G}$ by Lemma 1.32. Also, since $\mu_0(A) \le \mu_0(X) < \infty$ for all $A \in \mathcal{A}$ (see the comments after the proposition on the previous page), the function μ certainly maps into $[0,\infty)$.

Proof of (1). For any $A \in \mathcal{A}$ consider the sequence $(A_n)_{n\in\mathbb{N}}$ with $A_n = A$ for all n. Then $(A_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{A}$ such that $A_n \uparrow A$, and so $A \in \mathcal{G}$. Thus $\mathcal{A} \subseteq \mathcal{G}$, and since

$$\mu(A) = \lim_{n\to\infty} \mu_0(A_n) = \lim_{n\to\infty} \mu_0(A) = \mu_0(A),$$

it is also clear that $\mu|_{\mathcal{A}} = \mu_0$.

Proof of (2). Suppose $A, B \in \mathcal{G}$, so there exist sequences $(A_n)_{n\in\mathbb{N}}, (B_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ such that $A_n \uparrow A$ and $B_n \uparrow B$. Then it is clear that $A_n \cup B_n \uparrow A \cup B$ and $A_n \cap B_n \uparrow A \cap B$, where $A_n \cup B_n, A_n \cap B_n \in \mathcal{A}$ for all n since $\mathcal{A}$ is an algebra, and so $A \cup B, A \cap B \in \mathcal{G}$. Moreover, by Theorem 1.20(2),

$$\begin{aligned}
\mu(A \cup B) + \mu(A \cap B) &= \lim_{n\to\infty} \mu_0(A_n \cup B_n) + \lim_{n\to\infty} \mu_0(A_n \cap B_n) \\
&= \lim_{n\to\infty} \big(\mu_0(A_n \cup B_n) + \mu_0(A_n \cap B_n)\big) \\
&= \lim_{n\to\infty} \big(\mu_0(A_n) + \mu_0(B_n)\big) \\
&= \lim_{n\to\infty} \mu_0(A_n) + \lim_{n\to\infty} \mu_0(B_n) \\
&= \mu(A) + \mu(B).
\end{aligned}$$

where the usual limit laws apply by the remark following Lemma 1.32.

Proof of (3). Suppose $A, B \in \mathcal{G}$ with $B \subseteq A$. Let $(A_n)_{n\in\mathbb{N}}, (B_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ be such that $A_n \uparrow A$ and $B_n \uparrow B$. Then

$$\mu(B) = \lim_{n\to\infty} \mu_0(B_n) \le \lim_{n\to\infty} \mu_0(A_n) = \mu(A)$$

by Lemma 1.32.

Proof of (4). Suppose $(G_n)_{n\in\mathbb{N}} \subseteq \mathcal{G}$ with $G_n \uparrow G$. For each $m \in \mathbb{N}$ where is a sequence $(A_{m,n})_{n\in\mathbb{N}} \subseteq \mathcal{A}$ such that $A_{m,n} \uparrow_n G_m$. Define $(B_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ by

$$B_n = \bigcup_{m=1}^{n} A_{m,n},$$

where $B_n \subseteq B_{n+1}$ since $A_{m,n} \subseteq A_{m,n+1}$ for each $m, n \in \mathbb{N}$.

Let $x \in G$. Since

$$G = \bigcup_{m=1}^{\infty} G_m = \bigcup_{m=1}^{\infty}\bigcup_{n=1}^{\infty} A_{m,n},$$

$x \in A_{k,\ell}$ for some $k, \ell \in \mathbb{N}$. If $k \le \ell$, then $x \in B_\ell$ since $A_{k,\ell} \subseteq B_\ell$; and if $k > \ell$, then $x \in B_k$ since $A_{k,\ell} \subseteq A_{k,k} \subseteq B_k$. Thus $x \in \bigcup_{n=1}^\infty B_n$, and we see that $G \subseteq \bigcup_{n=1}^\infty B_n$. The reverse containment is clear, so $G = \bigcup_{n=1}^\infty B_n$ and we conclude that $B_n \uparrow G$. Therefore $G \in \mathcal{G}$, and moreover $\mu(G) = \lim_{n\to\infty} \mu_0(B_n)$.

Fix $m \in \mathbb{N}$. For all $n \ge m$ we have $A_{m,n} \subseteq B_n \subseteq G_n$, so that $\mu(A_{m,n}) \le \mu(B_n) \le \mu(G_n)$ by part (3), and then part (1) gives

$$\mu_0(A_{m,n}) \le \mu_0(B_n) \le \mu(G_n)$$

for all $n \ge m$. From this it follows that

$$\mu(G_m) = \lim_{n\to\infty} \mu_0(A_{m,n}) \le \lim_{n\to\infty} \mu_0(B_n) \le \lim_{n\to\infty} \mu(G_n),$$

and since $m \in \mathbb{N}$ is arbitrary we in turn obtain

$$\lim_{m\to\infty} \mu(G_m) \le \lim_{n\to\infty} \mu_0(B_n) \le \lim_{n\to\infty} \mu(G_n).$$

Therefore

$$\lim_{n\to\infty} \mu(G_n) = \lim_{n\to\infty} \mu_0(B_n) = \mu(G),$$

as was to be shown. ∎

Since $\mu(A) = \mu_0(A)$ for all $A \in \mathcal{A}$ and $\mathcal{A} \subseteq \mathcal{G}$, we see that μ is an extension of the premeasure μ_0 to a larger class of sets. Note that $\mathcal{G}$ is not necessarily an algebra, much less a σ-algebra.

Lemma 1.34. *Let collection $\mathcal{G} \subseteq \mathcal{P}(X)$ and set function $\mu : \mathcal{G} \to [0, \infty)$ have the following properties.*

P1. $\varnothing, X \in \mathcal{G}$, *with* $\mu(\varnothing) = 0$ *and* $\mu(X) = r$.

P2. *If* $G, H \in \mathcal{G}$, *then* $G \cup H, G \cap H \in \mathcal{G}$, *and*

$$\mu(G \cup H) + \mu(G \cap H) = \mu(G) + \mu(H).$$

P3. *If* $G, H \in \mathcal{G}$ *with* $H \subseteq G$, *then* $\mu(H) \le \mu(G)$.

P4. *If* $(G_n)_{n\in\mathbb{N}} \subseteq \mathcal{G}$ *with* $G_n \uparrow G$, *then* $G \in \mathcal{G}$ *and*

$$\lim_{n\to\infty} \mu(G_n) = \mu(G).$$

Define $\mu^* : \mathcal{P}(X) \to \overline{\mathbb{R}}$ *by*

$$\mu^*(A) = \inf\big\{\mu(G) : G \in \mathcal{G},\ G \supseteq A\big\}$$

for each $A \subseteq X$*. Then:*

Q1. $\mu^*|_{\mathcal{G}} = \mu$*, with* $0 \le \mu^*(A) \le r$ *for all* $A \in \mathcal{P}(X)$*.*

Q2. *For all* $A, B \in \mathcal{P}(X)$*,*

$$\mu^*(A \cup B) + \mu^*(A \cap B) \le \mu^*(A) + \mu^*(B),$$

and in particular $\mu^*(A) + \mu^*(A^c) \ge r$*.*

Q3. *If* $A_1, \dots, A_n \in \mathcal{P}(X)$*, then*

$$\mu^*\left(\bigcup_{k=1}^{n} A_k\right) \le \sum_{k=1}^{n} \mu^*(A_k).$$

Q4. *If* $A, B \in \mathcal{P}(X)$ *with* $B \subseteq A$*, then* $\mu^*(B) \le \mu^*(A)$*.*

Q5. *If* $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{P}(X)$ *with* $A_n \uparrow A$*, then*

$$\lim_{n\to\infty} \mu^*(A_n) = \mu^*(A).$$

Proof.

Proof of Q1. If $A \in \mathcal{G}$, then for all $G \in \mathcal{G}$ such that $G \supseteq A$ we have $\mu(G) \ge \mu(A)$ by P3, which shows $\mu(A)$ is a lower bound for the set $S = \{\mu(G) : G \in \mathcal{G},\ G \supseteq A\}$. Moreover, if $G' \in \mathcal{G}$ is such that $\mu(G') < \mu(A)$, then P3 makes clear that $G' \not\supseteq A$, and so $\mu(G') \notin S$. Hence $\mu(A)$ is the greatest lower bound for S; that is, $\mu^*(A) = \mu(A)$, and we conclude that $\mu^*|_{\mathcal{G}} = \mu$. Finally, since $\mu(X) = r$ and μ is nonnegative, by P3 it is clear that $\mu(G) \in [0, r]$ for all $G \in \mathcal{G}$, and hence $0 \le \mu^*(A) \le r$ for all $A \in \mathcal{P}(X)$.

Proof of Q2. Let $\epsilon > 0$. There exist $G, H \in \mathcal{G}$ such that $G \supseteq A$, $H \supseteq B$, and

$$\mu^*(A) \le \mu(G) < \mu^*(A) + \frac{\epsilon}{2} \quad \text{and} \quad \mu^*(B) \le \mu(H) < \mu^*(B) + \frac{\epsilon}{2}.$$

Now by P2, observing that $G \cup H \supseteq A \cup B$ and $G \cap H \supseteq A \cap B$,

$$\mu^*(A) + \mu^*(B) + \epsilon > \mu(G) + \mu(H) = \mu(G \cup H) + \mu(G \cap H) \ge \mu^*(A \cup B) + \mu^*(A \cap B).$$

Since $\epsilon > 0$ is arbitrary it follows that $\mu^*(A) + \mu^*(B) \ge \mu^*(A \cup B) + \mu^*(A \cap B)$. Finally,

$$\mu^*(A) + \mu^*(A^c) \ge \mu^*(A \cup A^c) + \mu^*(A \cap A^c) = \mu^*(X) + \mu^*(\varnothing) = \mu(X) + \mu(\varnothing) = r$$

by Q1 and P1.

Proof of Q3. For any $A, B \in \mathcal{P}(X)$ we have $\mu^*(A \cap B) \ge 0$ by Q1, and so

$$\mu^*(A \cup B) \le \mu^*(A) + \mu^*(B)$$

by Q2. The desired inequality now follows by induction.

Proof of Q4. Suppose $B \subseteq A$. Let $\epsilon > 0$. There exists some $G \in \mathcal{G}$ such that $G \supseteq A$ and $\mu^*(A) \leq \mu(G) < \mu^*(A) + \epsilon$. Now, since $G \supseteq B$ also, the definition of μ^* implies that $\mu(G) \geq \mu^*(B)$. Thus

$$\mu^*(A) + \epsilon > \mu(G) \geq \mu^*(B),$$

and since $\epsilon > 0$ is arbitrary it follows that $\mu^*(A) \geq \mu^*(B)$. ■

Theorem 1.35. *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra, let $\mu_0 : \mathcal{A} \to [0, \infty)$ be a finite premeasure with $\mu_0(X) = r$, and let*

$$\mathcal{G} = \{A \in \mathcal{P}(X) : \exists (A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}\,(A_n \uparrow A)\}.$$

Define $\mu : \mathcal{G} \to \overline{\mathbb{R}}$ by

$$\mu(A) = \lim_{n\to\infty} \mu_0(A_n),$$

where $(A_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{A}$ such that $A_n \uparrow A$. Further define $\mu^* : \mathcal{P}(X) \to \overline{\mathbb{R}}$ by

$$\mu^*(A) = \inf\{\mu(G) : G \in \mathcal{G},\ G \supseteq A\},$$

and let

$$\mathcal{H} = \{H \subseteq X : \mu^*(H) + \mu^*(H^c) = r\}.$$

Then $\mathcal{H}$ is a σ-algebra containing $\mathcal{G}$, and $\mu^* : \mathcal{H} \to [0, r]$ is a finite measure on $\mathcal{H}$. Therefore $(X, \mathcal{H}, \mu^*)$ is a finite measure space.

Proof. The function μ and collection $\mathcal{G}$ given here are the same μ and $\mathcal{G}$ in the hypothesis of Lemma 1.33, and thus they possess the properties (1)–(4) stated in that lemma's conclusion. These four properties easily imply the properties P1–P4 in the hypothesis of Lemma 1.34, and therefore μ^* possesses the properties Q1–Q5. In particular Q2 implies that

$$\mathcal{H} = \{H \subseteq X : \mu^*(H) + \mu^*(H^c) \leq r\}. \tag{1.6}$$

Lemmas 1.33 and 1.34 also give $\mu^*|_{\mathcal{G}} = \mu$ and $\mu|_{\mathcal{A}} = \mu_0$, and thus $\mu_0 = \mu = \mu^*$ on $\mathcal{A}$.

Let $G \in \mathcal{G}$, so there exists a sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ such that $A_n \uparrow G$, implying $G^c \subseteq A_n^c$ for all n. Now, $A_n^c \in \mathcal{G}$ since $\mathcal{A}$ is an algebra and $\mathcal{A} \subseteq \mathcal{G}$, so that

$$\mu_0(A_n) + \mu^*(G^c) \leq \mu_0(A_n) + \mu(A_n^c) = \mu_0(A_n) + \mu_0(A_n^c) = \mu_0(A_n \sqcup A_n^c) = \mu_0(X) = r,$$

recalling $\mu|_{\mathcal{A}} = \mu_0$ and the countable (hence finite) additivity of the premeasure μ_0. Since $\mu^*(A_n) = \mu_0(A_n)$, we have

$$\mu^*(A_n) + \mu^*(G^c) \leq r$$

for all n, so that by Q5

$$\mu^*(G) + \mu^*(G^c) = \lim_{n\to\infty} \mu^*(A_n) + \mu^*(G^c) \leq r,$$

implying $G \in \mathcal{H}$ and therefore $\mathcal{G} \subseteq \mathcal{H}$.

If $H \in \mathcal{H}$, then it is clear that $H^c \in \mathcal{H}$ as well. Also

$$\mu^*(X) + \mu^*(X^c) = \mu^*(X) + \mu^*(\varnothing) = \mu(X) + \mu(\varnothing) = r + 0 = r$$

by Q1 and P1, so $X \in \mathcal{H}$.

Suppose $H_1, H_2 \in \mathcal{H}$. By Q2,

$$\mu^*(H_1) + \mu^*(H_2) \geq \mu^*(H_1 \cup H_2) + \mu^*(H_1 \cap H_2), \tag{1.7}$$

and also

$$\mu^*(H_1^c) + \mu^*(H_2^c) \geq \mu^*(H_1^c \cup H_2^c) + \mu^*(H_1^c \cap H_2^c) = \mu^*(H_1 \cap H_2)^c + \mu^*(H_1 \cup H_2)^c, \tag{1.8}$$

Adding (1.7) and (1.8) yields

$$\begin{aligned} 2r &= \big[\mu^*(H_1) + \mu^*(H_1^c)\big] + \big[\mu^*(H_2) + \mu^*(H_2^c)\big] \\ &\geq \big[\mu^*(H_1 \cup H_2) + \mu^*(H_1 \cup H_2)^c\big] + \big[\mu^*(H_1 \cap H_2) + \mu^*(H_1 \cap H_2)^c\big] \geq 2r, \end{aligned}$$

the second inequality owing to Q2. Thus

$$\big[\mu^*(H_1 \cup H_2) + \mu^*(H_1 \cup H_2)^c\big] + \big[\mu^*(H_1 \cap H_2) + \mu^*(H_1 \cap H_2)^c\big] = 2r,$$

implying that

$$\big[\mu^*(H_1 \cup H_2) + \mu^*(H_1 \cup H_2)^c\big] \leq r,$$

whereupon Q2 leads to the conclusion that

$$\big[\mu^*(H_1 \cup H_2) + \mu^*(H_1 \cup H_2)^c\big] = r,$$

Therefore $H_1 \cup H_2 \in \mathcal{H}$, which implies that $\mathcal{H}$ is closed under finite unionization and so is an algebra.

Next, let $(H_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $H_n \uparrow H$. Let $\epsilon > 0$. Now, $\lim_{n\to\infty} \mu^*(H_n) = \mu^*(H)$ by Q5, and so there exists some $m \in \mathbb{N}$ such that $\mu^*(H) \leq \mu^*(H_n) + \epsilon$ for all $n \geq m$. Also, $H^c \subseteq H_n^c$ for all n, and so $\mu^*(H^c) \leq \mu^*(H_n^c)$ for all n by Q4. Hence we have

$$\mu^*(H) + \mu^*(H^c) \leq \mu^*(H_n) + \mu^*(H_n^c) + \epsilon = r + \epsilon$$

for all $n \geq m$, and since $\epsilon > 0$ is arbitrary, it follows that $\mu^*(H) + \mu^*(H^c) \leq r$, and therefore $H \in \mathcal{H}$. Now, if $(H_n)_{n\in\mathbb{N}}$ is an arbitrary sequence in $\mathcal{H}$, observe that $(\bigcup_{k=1}^n H_k)_{n\in\mathbb{N}}$ is also a sequence in $\mathcal{H}$ since $\mathcal{H}$ is closed under finite unionization, and since

$$\bigcup_{k=1}^n H_k \uparrow \bigcup_{k=1}^\infty H_k,$$

we conclude that $\bigcup_{k=1}^\infty H_k \in \mathcal{H}$. So $\mathcal{H}$ is closed under countable unionization, and therefore $\mathcal{H}$ is a σ-algebra.

Finally, (1.7) must in fact be an equality, for otherwise adding (1.7) and (1.8) would yield the contradiction $2r > 2r$. Thus

$$\mu^*(H_1 \cup H_2) + \mu^*(H_1 \cap H_2) = \mu^*(H_1) + \mu^*(H_2)$$

for all $H_1, H_2 \in \mathcal{H}$, and so if H_1 and H_2 are disjoint we obtain $\mu^*(H_1 \cup H_2) = \mu^*(H_1) + \mu^*(H_2)$. This implies, by induction, that μ^* is a finitely additive set function on the σ-algebra $\mathcal{H}$. Moreover, given any $H \in \mathcal{H}$ and sequence $(H_n)_{n\in\mathbb{N}} \subseteq \mathcal{H}$ such that $H_n \uparrow H$, we have $\lim_{n\to\infty} \mu^*(H_n) = \mu^*(H)$ by Q5. So μ^* is continuous from below on $\mathcal{H}$, and hence countably additive on $\mathcal{H}$ by Theorem 1.23. Since $\mu^*(\varnothing) = 0$, and $\mu^*(H) \in [0, r]$ for all $H \in \mathcal{H}$ by Q1, we have now shown that $\mu^* : \mathcal{H} \to [0, r]$ is a finite measure on $\mathcal{H}$. ∎

Corollary 1.36. *Any finite premeasure on an algebra $\mathcal{A}$ can be extended to a finite measure on $\sigma(\mathcal{A})$.*

Proof. If μ_0 is a finite premeasure on an algebra $\mathcal{A}$, then Theorem 1.35 shows that μ_0 can be extended to a finite measure μ^* on a σ-algebra $\mathcal{H}$ containing $\mathcal{A}$. Since $\mathcal{A} \subseteq \sigma(\mathcal{A}) \subseteq \mathcal{H}$, it follows immediately that μ^* restricted to $\sigma(\mathcal{A})$ is an extension of μ_0 to a finite measure on $\sigma(\mathcal{A})$. ■

If μ is a measure on a σ-algebra $\mathcal{S}$, then a **μ-null set** (or simply **null set** if there is no chance of ambiguity) is a set $A \in \mathcal{S}$ such that $\mu(A) = 0$. See denote by $\mathcal{N}_\mu$ the collection of all μ-null sets:

$$\mathcal{N}_\mu = \{A \in \mathcal{S} : \mu(A) = 0\}.$$

It will be convenient to define

$$\mathcal{P}(\mathcal{N}_\mu) = \bigcup_{N \in \mathcal{N}_\mu} \mathcal{P}(N),$$

which is the collection of all subsets of μ-null sets. Not all such sets are necessarily in $\mathcal{S}$, but it is often desirable that a measure space $(X, \mathcal{S}, \mu)$ be such that $\mathcal{P}(\mathcal{N}_\mu) \subseteq \mathcal{S}$.

Definition 1.37. *A measure space $(X, \mathcal{S}, \mu)$ is **complete** if $\mathcal{P}(\mathcal{N}_\mu) \subseteq \mathcal{S}$, in which case we say μ is **complete on** $\mathcal{S}$.*

Proposition 1.38. *The measure space $(X, \mathcal{H}, \mu^*)$ constructed in Theorem 1.35 is complete.*

Proof. Let $A \in \mathcal{H}$ be such that $\mu^*(A) = 0$, so $r = \mu^*(A) + \mu^*(A^c) = \mu^*(A^c)$. Let $B \subseteq A$. By Q1 and Q4, $0 \le \mu^*(B) \le \mu^*(A) = 0$, so $\mu^*(B) = 0$. Since $B^c \supseteq A^c$, $r = \mu^*(A^c) \le \mu^*(B^c) \le r$ by Q1 and Q4 again, so $\mu^*(B^c) = r$. Hence $\mu^*(B) + \mu^*(B^c) = r$, implying that $B \in \mathcal{H}$ and therefore μ is complete on $\mathcal{H}$. ■

Remark. A quicker way to prove Proposition 1.38 is to note that

$$\mu^*(B) + \mu^*(B^c) \le \mu^*(A) + \mu^*(B^c) = \mu^*(B^c) \le r$$

and then use (1.6).

Theorem 1.39. *Let $(X, \mathcal{S}, \mu)$ be a measure space, let*

$$\mathcal{S}_\mu = \big\{A \cup N : A \in \mathcal{S},\ N \in \mathcal{P}(\mathcal{N}_\mu)\big\},$$

and define $\bar{\mu} : \mathcal{S}_\mu \to \overline{\mathbb{R}}$ by

$$\bar{\mu}(A \cup N) = \mu(A)$$

for all $A \cup N \in \mathcal{S}_\mu$. Then $(X, \mathcal{S}_\mu, \bar{\mu})$ is a complete measure space.

Proof. It is first necessary to verify that $\mathcal{S}_\mu$ is a σ-algebra. Since $\varnothing \in \mathcal{N}_\mu \subseteq \mathcal{P}(\mathcal{N}_\mu)$, it is clear that $\mathcal{S} \subseteq \mathcal{S}_\mu$, and hence $X \in \mathcal{S}_\mu$ in particular. Also $\mathcal{P}(\mathcal{N}_\mu) \subseteq \mathcal{S}_\mu$, since any $N \in \mathcal{P}(\mathcal{N}_\mu)$ can be written as $N = \varnothing \cup N$ for $\varnothing \in \mathcal{S}$.

Let $(S_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{S}_\mu$, so $S_n \in A_n \cup N_n$ for some $A_n \in \mathcal{S}$ and $N_n \in \mathcal{P}(\mathcal{N}_\mu)$. For each n there is some $M_n \in \mathcal{N}_\mu$ such that $N_n \subseteq M_n$. Now,

$$\mu\left(\bigcup_{n=1}^{\infty} M_n\right) = \sum_{n=1}^{\infty} \mu(M_n) = 0$$

by the countable additivity of μ on $\mathcal{S}$, which implies that $\bigcup_{n=1}^\infty M_n \in \mathcal{N}_\mu$, and since $\bigcup_{n=1}^\infty N_n \subseteq \bigcup_{n=1}^\infty M_n$ we see that $\bigcup_{n=1}^\infty N_n \in \mathcal{P}(\mathcal{N}_\mu)$. So

$$\bigcup_{n=1}^\infty S_n = \left(\bigcup_{n=1}^\infty A_n\right) \cup \left(\bigcup_{n=1}^\infty N_n\right)$$

with $\bigcup_{n=1}^\infty A_n \in \mathcal{S}$ and $\bigcup_{n=1}^\infty N_n \in \mathcal{P}(\mathcal{N}_\mu)$, which shows that $\bigcup_{n=1}^\infty S_n \in \mathcal{S}_\mu$. Therefore $\mathcal{S}_\mu$ is closed under countable unionization.

Next, suppose $S \in \mathcal{S}_\mu$, so $S = A \cup N$ for some $A \in \mathcal{S}$ and $N \in \mathcal{P}(\mathcal{N}_\mu)$. There exists some $M \in \mathcal{N}_\mu$ such that $N \subseteq M$, where

$$(A \cup N)^c = A^c \cap N^c = A^c \cap \big((M \setminus N) \cup M^c\big) = \big(A^c \cap (M \setminus N)\big) \cup (A^c \cap M^c).$$

Now, $A^c \cap (M \setminus N) \subseteq M$ implies $A^c \cap (M \setminus N) \in \mathcal{P}(\mathcal{N}_\mu) \subseteq \mathcal{S}_\mu$, whereas $A^c \cap M^c \in \mathcal{S} \subseteq \mathcal{S}_\mu$; and so, since $\mathcal{S}_\mu$ is closed under finite unionization, it follows that $(A \cup N)^c \in \mathcal{S}_\mu$. That is, $\mathcal{S}_\mu$ is closed under complementation and is therefore a σ-algebra.

Next we verify that $\bar{\mu}$ is a measure on $\mathcal{S}_\mu$. Clearly $\bar{\mu}(\varnothing) = 0$ and $\bar{\mu}(S) \geq 0$ for all $S \in \mathcal{S}_\mu$. As for countable additivity, referencing the sequence $(S_n)_{n\in\mathbb{N}} \subseteq \mathcal{S}_\mu$ with $S_n = A_n \cup N_n$ above, only now assuming the sequence to be disjoint, we have

$$\begin{aligned}\bar{\mu}\left(\bigcup_{n=1}^\infty S_n\right) &= \bar{\mu}\left(\left(\bigcup_{n=1}^\infty A_n\right) \cup \left(\bigcup_{n=1}^\infty N_n\right)\right) = \mu\left(\bigcup_{n=1}^\infty A_n\right) \\ &= \sum_{n=1}^\infty \mu(A_n) = \sum_{n=1}^\infty \bar{\mu}(A_n \cup N_n) = \sum_{n=1}^\infty \bar{\mu}(S_n),\end{aligned}$$

the middle equality justified since the sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{S}$ is also disjoint. Thus $\bar{\mu} : \mathcal{S}_\mu \to \overline{\mathbb{R}}$ is a measure, and $(X, \mathcal{S}_\mu, \bar{\mu})$ is a measure space.

Suppose $S \in \mathcal{P}(\mathcal{N}_{\bar{\mu}})$. Then $S \subseteq M$ for some $M \in \mathcal{S}_\mu$ such that $\bar{\mu}(M) = 0$. Now, $M \in \mathcal{S}_\mu$ implies $M = A \cup N$ for some $A \in \mathcal{S}$ and $N \in \mathcal{P}(\mathcal{N}_\mu)$, which in turn implies there exists some $N' \in \mathcal{N}_\mu$ such that $M \subseteq A \cup N' \in \mathcal{S}$. Since

$$0 = \bar{\mu}(M) = \bar{\mu}(A \cup N) = \bar{\mu}(A) + \bar{\mu}(N) \quad \Rightarrow \quad \mu(A) = \bar{\mu}(A) = 0,$$

we see that $\mu(A \cup N') = \mu(A) + \mu(N') = 0$, and hence $A \cup N' \in \mathcal{N}_\mu$. Since $S \subseteq M \subseteq A \cup N'$, it follows that $S \in \mathcal{P}(\mathcal{N}_\mu) \subseteq \mathcal{S}_\mu$, and therefore $\mathcal{P}(\mathcal{N}_{\bar{\mu}}) \subseteq \mathcal{S}_\mu$. We conclude that $(X, \mathcal{S}_\mu, \bar{\mu})$ is a complete measure space. ■

Definition 1.40. *The measure space $(X, \mathcal{S}_\mu, \bar{\mu})$ in Theorem 1.39 is the **completion** of $(X, \mathcal{S}, \mu)$, and we say $\mathcal{S}_\mu$ is the **completion of** $\mathcal{S}$ **relative to** μ.*

Proposition 1.41. *The measure space $(X, \mathcal{H}, \mu^*)$ constructed in Theorem 1.35 is the completion of $(X, \sigma(\mathcal{A}), \mu^*)$.*

A class $\mathcal{C} \subseteq \mathcal{P}(X)$ is **monotone** if whenever a sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}$ is such that $A_n \uparrow A$ or $A_n \downarrow A$, then $A \in \mathcal{C}$.

Theorem 1.42 (Monotone Class Theorem). *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra, and let $\mathcal{C} \subseteq \mathcal{P}(X)$ be monotone. If $\mathcal{A} \subseteq \mathcal{C}$, then $\sigma(\mathcal{A}) \subseteq \mathcal{C}$.*

Proof. Let $\mathcal{M} \subseteq \mathcal{P}(X)$ be the smallest monotone class such that $\mathcal{A} \subseteq \mathcal{M}$. We shall show that $\mathcal{M}$ is in fact a σ-algebra.

Fix $F \in \mathcal{A}$, and define

$$\mathcal{M}_F = \left\{ M \in \mathcal{M} : F \cap M, F \cap M^c, F^c \cap M \in \mathcal{M} \right\}.$$

Let $(M_n)_{n\in\mathbb{N}} \subseteq \mathcal{M}_F$ be such that $M_n \uparrow\downarrow A$, which here we take to mean that either $M_n \uparrow A$ or $M_n \downarrow A$. Since $(M_n)_{n\in\mathbb{N}} \subseteq \mathcal{M}$ it is immediate that $A \in \mathcal{M}$. Moreover, $F \cap M_n$, $F \cap M_n^c$, and $F^c \cap M_n$ are all elements of $\mathcal{M}$ for each n, and since $\mathcal{M}$ is a monotone class and

$$F \cap M_n \uparrow\downarrow F \cap A, \quad F \cap M_n^c \downarrow\uparrow F \cap A^c, \quad F^c \cap M_n \uparrow\downarrow F^c \cap A,$$

it follows that $F \cap A, F \cap A^c, F^c \cap A \in \mathcal{M}$. Hence $A \in \mathcal{M}_F$, and we conclude that $\mathcal{M}_F$ is a monotone class. Clearly $\mathcal{A} \subseteq \mathcal{M}_F \subseteq \mathcal{M}$, and since $\mathcal{M}$ is the smallest monotone class containing $\mathcal{A}$, we see that $\mathcal{M}_F = \mathcal{M}$ for all $F \in \mathcal{A}$.

Next let $B \in \mathcal{M}$ be arbitrary. Then $B \in \mathcal{M}_F$ for any $F \in \mathcal{A}$, which is to say $F \cap B$, $F \cap B^c$, and $F^c \cap B$ are all elements of $\mathcal{M}$ for any $F \in \mathcal{A}$, and thus

$$\mathcal{A} \subseteq \left\{ M \in \mathcal{M} : B \cap M, B \cap M^c, B^c \cap M \in \mathcal{M} \right\} = \mathcal{M}_B.$$

Now, $\mathcal{A} \subseteq \mathcal{M}_B \subseteq \mathcal{M}$, and since $\mathcal{M}_B$ is a monotone class (by the same argument that showed $\mathcal{M}_F$ is a monotone class), it follows that $\mathcal{M}_B = \mathcal{M}$ for any $B \in \mathcal{M}$.

Let $A, B \in \mathcal{M}$. Since $\mathcal{M} = \mathcal{M}_B$, we have $B \cap A, B \cap A^c, B^c \cap A \in \mathcal{M}$. Since $\mathcal{A} \subseteq \mathcal{M}$ we clearly have $X \in \mathcal{M}$, and then $A^c = X \cap A^c \in \mathcal{M}$ shows $\mathcal{M}$ is closed under complementation. So $A^c, B^c \in \mathcal{M}$, and since $\mathcal{M}$ is also closed under finite intersection,

$$A^c, B^c \in \mathcal{M} \;\Rightarrow\; A^c \cap B^c \in \mathcal{M} \;\Rightarrow\; (A \cup B)^c \in \mathcal{M} \;\Rightarrow\; A \cup B \in \mathcal{M},$$

and we find that $\mathcal{M}$ is closed under finite unionization. Thus $\mathcal{M}$ is an algebra as well as a monotone class, which easily implies that $\mathcal{M}$ is a σ-algebra; for if $(M_n)_{n\in\mathbb{N}} \subseteq \mathcal{M}$, we have $\bigcup_{k=1}^n M_k \in \mathcal{M}$ for each n with

$$\bigcup_{k=1}^{n} M_k \uparrow \bigcup_{k=1}^{\infty} M_k,$$

and therefore $\bigcup_{n=1}^\infty M_n \in \mathcal{M}$.

Since $\mathcal{M}$ is a σ-algebra containing $\mathcal{A}$, we have $\mathcal{M} \supseteq \sigma(\mathcal{A})$. But $\mathcal{M}$ is also the smallest monotone class containing $\mathcal{A}$, so that $\mathcal{M} \subseteq \mathcal{C}$. Therefore $\sigma(\mathcal{A}) \subseteq \mathcal{C}$ and we are done. ∎

Theorem 1.43 (Carathéodory Extension Theorem). *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra. If μ is a σ-finite premeasure on $\mathcal{A}$, then μ has a unique extension to a measure $\mu^* : \sigma(\mathcal{A}) \to \overline{\mathbb{R}}$.*

Theorem 1.44 (Approximation Theorem). *Let $(X, \mathcal{S}, \mu)$ be a measure space, and let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra such that $\sigma(\mathcal{A}) = \mathcal{S}$. Assume $\mu|_{\mathcal{A}}$ is a σ-finite premeasure, and let $\epsilon > 0$. If $A \in \mathcal{S}$ and $\mu(A) < \infty$, then there is some $B \in \mathcal{A}$ such that $\mu(A \Delta B) < \epsilon$.*

1.5 – Outer Measures

Definition 1.45. *An **outer measure** on X is a set function $\mu^* : \mathcal{P}(X) \to [0,\infty]$ with the following properties.*

OM1. $\mu^*(\varnothing) = 0$.

OM2. *Monotonicity: if $A, B \in \mathcal{P}(X)$ with $B \subseteq A$, then $\mu^*(B) \le \mu^*(A)$.*

OM3. *Countable Subadditivity: if $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{P}(X)$, then*

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

*A set $E \subseteq X$ is **μ^*-measurable** if*

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

for all $A \subseteq X$.

Proposition 1.46. *The set function μ^* in Lemma 1.34 (and hence also in Theorem 1.35) is an outer measure on X.*

Proof. Since μ maps into $[0,\infty)$, it is clear that μ^* maps into $[0,\infty]$. Property OM1 derives from P1 together with Q1 in Lemma 1.34, and OM2 derives from Q4. Finally, if $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{P}(X)$ then $\bigcup_{k=1}^{n} A_k \uparrow \bigcup_{n=1}^{\infty} A_n$, so by Q5 followed by Q4 we have

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu^*\left(\bigcup_{k=1}^{n} A_k\right) \le \lim_{n\to\infty} \sum_{k=1}^{n} \mu^*(A_k) = \sum_{n=1}^{\infty} \mu^*(A_n),$$

securing OM3. ∎

Theorem 1.47 (Carathéodory Restriction Theorem). *Let ρ be an outer measure on X. If $\mathcal{M}$ is the class of all ρ-measurable sets, then $(X, \mathcal{M}, \rho)$ is a complete measure space.*

Proof. For all $A \in \mathcal{P}(X)$,

$$\rho(A \cap X) + \rho(A \cap X^c) = \rho(A) + \rho(\varnothing) = \rho,$$

so $X \in \mathcal{M}$. It is also clear that $E \in \mathcal{M}$ implies $E^c \in \mathcal{M}$. Let $E, F \in \mathcal{M}$, so

$$\rho(A) = \rho(A \cap E) + \rho(A \cap E^c) \quad \text{and} \quad \rho(A) = \rho(A \cap F) + \rho(A \cap F^c)$$

for all $A \in \mathcal{P}(X)$. Let $E, F \in \mathcal{M}$. For any $A \in \mathcal{P}(X)$, by the subadditivity (OM3) and monotonicity (OM2) properties of ρ,

$$\begin{aligned}
&\rho\big(A \cap (E \cup F)\big) + \rho\big(A \cap (E \cup F)^c\big) \\
&\qquad \le \rho(A \cap E) + \rho(A \cap F) + \rho\big((A \cap E^c) \cap F^c\big) \\
&\qquad = \big[\rho(A) - \rho(A \cap E^c)\big] + \big[\rho(A) - \rho(A \cap F^c)\big] + \big[\rho(A \cap E^c) - \rho(A \cap E^c \cap F)\big] \\
&\qquad = 2\rho(A) - \big[\rho(A \cap F^c) + \rho\big((A \cap E^c) \cap F\big)\big]
\end{aligned}$$

$$\leq 2\rho(A) - \big[\rho\big((A \cap E^c) \cap F^c\big) + \rho\big((A \cap E^c) \cap F\big)\big]$$
$$= 2\rho(A) - \rho(A) = \rho(A).$$

By subadditivity

$$\rho\big(A \cap (E \cup F)\big) + \rho\big(A \cap (E \cup F)^c\big) \geq \rho(A),$$

and so

$$\rho\big(A \cap (E \cup F)\big) + \rho\big(A \cap (E \cup F)^c\big) = \rho(A).$$

Thus $E \cup F \in \mathcal{M}$, and it follows by induction that $\mathcal{M}$ is closed under finite unionization. Therefore $\mathcal{M}$ is an algebra, and moreover if E and F are disjoint we have

$$\rho\big(A \cap (E \cup F)\big) = \rho\Big(\big(A \cap (E \cup F)\big) \cap E\Big) + \rho\Big(\big(A \cap (E \cup F)\big) \cap E^c\Big)$$
$$= \rho(A \cap E) + \rho(A \cap F)$$

for any $A \in \mathcal{P}(X)$, and so by induction we obtain

$$\rho\left(A \cap \bigsqcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \rho(A \cap E_k) \tag{1.9}$$

for any finite disjoint sequence $(E_k)_{k=1}^n \subseteq \mathcal{M}$.

Let $(E_n)_{n\in\mathbb{N}} \subseteq \mathcal{M}$ be a disjoint sequence, and let $E = \bigsqcup_{n=1}^{\infty} E_n$. Fix $A \in \mathcal{P}(X)$. For each n we have $F_n = \bigcup_{k=1}^{n} E_k \in \mathcal{M}$, where

$$F_n^c = \bigcap_{k=1}^{n} E_k^c \supseteq \bigcap_{k=1}^{\infty} E_k^c = E^c,$$

so that

$$\rho(A) = \rho(A \cap F_n) + \rho(A \cap F_n^c) \geq \rho(A \cap F_n) + \rho(A \cap E^c)$$

by OM2, and thus

$$\rho(A) \geq \sum_{k=1}^{n} \rho(A \cap E_k) + \rho(A \cap E^c)$$

by (1.9). Letting $n \to \infty$ then gives

$$\rho(A) \geq \sum_{n=1}^{\infty} \rho(A \cap E_n) + \rho(A \cap E^c)$$

for all $A \in \mathcal{P}(X)$. By OM2,

$$\rho(A) \geq \rho\left(\bigcup_{n=1}^{\infty} (A \cap E_n)\right) + \rho(A \cap E^c) = \rho(A \cap E) + \rho(A \cap E^c).$$

This, together with OM3, yields $\rho(A) = \rho(A \cap E) + \rho(A \cap E^c)$, and thus $E \in \mathcal{M}$. Now, if $(E_n)_{n\in\mathbb{N}} \subseteq \mathcal{M}$ is any arbitrary sequence (not necessarily disjoint), by Theorem 1.4 we find that

$$E := \bigcup_{n=1}^{\infty} E_n = \bigsqcup_{n=1}^{\infty} (E_1^c \cap \cdots \cap E_{n-1}^c \cap E_n),$$

where the disjoint sequence of sets $E_1^c \cap \cdots \cap E_{n-1}^c \cap E_n$ is a sequence in $\mathcal{M}$ since $\mathcal{M}$ is an algebra, and so we find that $E \in \mathcal{M}$ once more. Therefore $\mathcal{M}$ is a σ-algebra.

For the next step, assume again that $(E_n)_{n\in\mathbb{N}} \subseteq \mathcal{M}$ is a disjoint sequence. Letting $A = \bigsqcup_{k=1}^{\infty} E_k$ in (1.9), we have

$$\rho\left(\bigsqcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \rho(E_k)$$

for all n, and so

$$\lim_{n\to\infty} \rho\left(\bigsqcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{\infty} \rho(E_k).$$

Now, by OM2 and OM3,

$$\rho\left(\bigsqcup_{k=1}^{n} E_k\right) \leq \rho\left(\bigsqcup_{k=1}^{\infty} E_k\right) \leq \sum_{k=1}^{\infty} \rho(E_k)$$

for all n, implying that

$$\sum_{k=1}^{\infty} \rho(E_k) = \lim_{n\to\infty} \rho\left(\bigsqcup_{k=1}^{n} E_k\right) \leq \rho\left(\bigsqcup_{k=1}^{\infty} E_k\right) \leq \sum_{k=1}^{\infty} \rho(E_k),$$

and hence

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

Therefore ρ is a measure on $\mathcal{M}$, and so $(X, \mathcal{M}, \rho)$ is a measure space.

Finally, suppose $E \in \mathcal{P}(X)$ is such that $\rho(E) = 0$. Let $A \in \mathcal{P}(X)$ be arbitrary. From

$$0 \leq \rho(A \cap E) \leq \rho(E) = 0$$

we have $\rho(A \cap E) = 0$, and then

$$\rho(A) = \rho\big((A \cap E) \cup (A \cap E^c)\big) \leq \rho(A \cap E) + \rho(A \cap E^c) = \rho(A \cap E^c) \leq \rho(A)$$

implies that $\rho(A \cap E) + \rho(A \cap E^c) = \rho(A)$. Thus E is ρ-measurable, and we conclude that $E \in \mathcal{M}$ for any $E \subseteq X$ with $\rho(E) = 0$. Now, if $A \subseteq E$ for some $E \in \mathcal{M}$ with $\rho(E) = 0$, we find that $A \in \mathcal{M}$ since $0 \leq \rho(A) \leq \rho(E) = 0$ implies $\rho(A) = 0$. Therefore $(X, \mathcal{M}, \rho)$ is complete. ■

Proposition 1.48. *Let μ be a premeasure on an algebra $\mathcal{A} \subseteq \mathcal{P}(X)$. Define $\mu^* : \mathcal{P}(X) \to \overline{\mathbb{R}}$ by*

$$\mu^*(A) = \inf\left\{\sum_n \mu(E_n) : A \subseteq \bigcup_n E_n \text{ for } E_n \in \mathcal{A}\right\}.$$

Then μ^* is an outer measure on X such that $\mu^*|_{\mathcal{A}} = \mu$.

Proof. Since $\mu : \mathcal{A} \to [0, \infty]$, it is clear that $\mu^* : \mathcal{P}(X) \to [0, \infty]$. Also it is clear that $\mu^*(\varnothing) = 0$, satisfying OM1.

For each $\Omega \in \mathcal{P}(X)$ define

$$S_\Omega = \left\{\sum_n \mu(E_n) : \Omega \subseteq \bigcup_n E_n \text{ for } E_n \in \mathcal{A}\right\}.$$

Suppose $A, B \in \mathcal{P}(X)$ with $B \subseteq A$. Let $s \in S_A$, so $s = \sum_n \mu(E_n)$ for some $E_n \in \mathcal{A}$ such that $A \subseteq \bigcup_n E_n$. Since $B \subseteq \bigcup_n E_n$ also, it follows that $s \in S_B$, and so $S_A \subseteq S_B$. It follows immediately that

$$\mu^*(B) = \inf(S_B) \leq \inf(S_A) = \mu^*(A),$$

and so μ^* satisfies OM2.

Next, let $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{P}(X)$. Fix $\epsilon > 0$. For each n there exist $(E_{nk})_{k\in\mathbb{N}} \subseteq \mathcal{A}$ such that

$$A_n \subseteq \bigcup_{k=1}^{\infty} E_{nk} \quad \text{and} \quad \sum_{k=1}^{\infty} \mu(E_{nk}) \leq \frac{\epsilon}{2^n} + \mu^*(A_n).$$

(Each sequence $(E_{nk})_{k\in\mathbb{N}}$ may be assumed to be infinite by taking $E_{nk} = \varnothing$ for all sufficiently large k.) Now,

$$\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} E_{nk},$$

and so by definition of μ^* we have

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mu(E_{nk}) \leq \sum_{n=1}^{\infty} \left[\mu^*(A_n) + \epsilon 2^{-n}\right] \leq \sum_{n=1}^{\infty} \mu^*(A_n) + \epsilon.$$

Since $\epsilon > 0$ is arbitrary we conclude that

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

and so μ^* satisfies OM3 and is therefore an outer measure on X. That $\mu^*|_{\mathcal{A}} = \mu$ is easily verified. ■

1.6 – Lebesgue-Stieltjes Measures and Distribution Functions

Given any metric space (Y, d), a function $f : A \subseteq \mathbb{R} \to Y$ is **right-continuous** at $a \in A$ if $\lim_{x\to a^+} f(x) = f(a)$, and **left-continuous** at $a \in A$ if $\lim_{x\to a^-} f(x) = f(a)$.

Definition 1.49. *A **Lebesgue-Stieltjes measure** on $\mathbb{R}$ is a measure $\mu : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ such that $\mu(I) < \infty$ for every bounded interval I. A **distribution function** on $\mathbb{R}$ is a monotone increasing right-continuous function $F : \mathbb{R} \to \mathbb{R}$.*

Theorem 1.50. *Let μ be a Lebesgue-Stieltjes measure on $\mathbb{R}$. If $F : \mathbb{R} \to \mathbb{R}$ is given by*

$$F(b) - F(a) = \mu(a, b] \tag{1.10}$$

for all $-\infty < a < b < \infty$, then F is a distribution function.

The equation (1.10) only determines F up to an arbitrary constant. One way to see this is to note that

$$F(x) = \begin{cases} \mu(x, 0] + F(0), & x < 0 \\ F(0), & x = 0 \\ \mu(0, x] + F(0), & x > 0, \end{cases}$$

which shows that F may be uniquely determined by choosing the value of $F(0)$.

Given a distribution function $F : \mathbb{R} \to \mathbb{R}$, we can extend the definition of F to $\overline{\mathbb{R}}$ by setting

$$F(\infty) = \lim_{x\to\infty} F(x) \quad \text{and} \quad F(-\infty) = \lim_{x\to-\infty} F(x),$$

where both limits exist as extended real numbers since F is monotone increasing. Now define $\mu(a, b] = F(b) - F(a)$ for all $a, b \in \overline{\mathbb{R}}$, and in particular

$$\mu[-\infty, b] = F(b) - F(-\infty) = \mu(-\infty, b].$$

This defines μ on the collection $\mathcal{H}_{\overline{\mathbb{R}}}$ of all h-intervals of $\overline{\mathbb{R}}$, which by convention we take to include intervals of the form $[-\infty, b]$. If $(I_k)_{k=1}^n$ is a disjoint finite sequence in $\mathcal{H}_{\overline{\mathbb{R}}}$, define

$$\mu\left(\bigsqcup_{k=1}^n I_k\right) = \sum_{k=1}^n \mu(I_k).$$

This extends μ to the algebra $\mathcal{A}(\overline{\mathbb{R}})$ of finite disjoint unions of h-intervals of $\overline{\mathbb{R}}$. Since $\overline{\mathbb{R}}$ is compact, it can be shown that μ is in fact countably additive on $\mathcal{A}(\overline{\mathbb{R}})$. Thus, since $\mu(\varnothing) = 0$ by Theorem 1.20(1), it follows that $(\mathbb{R}, \mathcal{A}(\overline{\mathbb{R}}), \mu)$ is a premeasure space. It then can be shown that $(\mathbb{R}, \mathcal{A}(\mathbb{R}), \mu)$ is a premeasure space, where $\mathcal{A}(\mathbb{R})$ is the algebra of finite disjoint unions of h-intervals of $\mathbb{R}$, which by convention we take to include intervals of the form (a, ∞). The Carathéodory Extension Theorem then gives a unique extension of μ to a measure μ^* to $\sigma(\mathcal{A}(\mathbb{R})) = \mathcal{B}(\mathbb{R})$. We state the theorem.

Theorem 1.51. *Let F be a distribution function on $\mathbb{R}$. If $\mu(a, b] = F(b) - F(a)$ for all $-\infty < a < b < \infty$, then there is a unique extension of μ to a Lebesgue-Stieltjes measure on $\mathbb{R}$.*

Theorems 1.50 and 1.51 together demonstrate that there is a one-to-one correspondence between Lebesgue-Stieltjes measure on $\mathbb{R}$ and distribution functions (up to an additive constant) on $\mathbb{R}$. The Lebesgue-Stieltjes measure corresponding to a distribution function F is called the **measure induced by** F. The distribution function F (up to an additive constant) corresponding to a Lebesgue-Stieltjes measure μ is called the **distribution function of** μ.

From the formulation $\mu(a,b] = F(b) - F(a)$ for $-\infty < a < b < \infty$ given in Theorem 1.51 we may express intervals that are not bounded h-intervals of $\mathbb{R}$ in terms of F. Defining

$$F(c^-) = \lim_{x \to c^-} F(x)$$

for any $c \in \mathbb{R}$, the following can be proved:

$$\mu(a,b) = F(b^-) - F(a), \quad \mu[a,b] = F(b) - F(a^-), \quad \mu[a,b) = F(b^-) - F(a^-), \tag{1.11}$$

$$\mu(-\infty, x] = F(x) - F(-\infty), \quad \mu(-\infty, x) = F(x^-) - F(-\infty), \tag{1.12}$$

$$\mu(x, \infty) = F(\infty) - F(x), \quad \mu[x, \infty) = F(\infty) - F(x^-), \tag{1.13}$$

and

$$\mu(\mathbb{R}) = F(\infty) - F(-\infty). \tag{1.14}$$

A special case obtaining from $\mu[a,b] = F(b) - F(a^-)$ is

$$\mu\{x\} = F(x) - F(x^-), \tag{1.15}$$

which shows that a distribution function F corresponding to a measure μ is continuous at $x \in \mathbb{R}$ if and only if $\mu\{x\} = 0$.

Example 1.52. Recall the Riemann integral $\int_a^b f$ of a function f on a closed bounded interval $[a,b]$.[2] Let $f : \mathbb{R} \to [0, \infty)$ be such that f is Riemann integrable on any closed bounded interval, and let $F(0) \in \mathbb{R}$ be arbitrary. Define

$$F(x) - F(0) = \int_0^x f$$

for all $x \in \mathbb{R} \setminus \{0\}$, where as usual

$$\int_0^x f := -\int_x^0 f$$

if $x < 0$. Then $F : \mathbb{R} \to \mathbb{R}$ is a continuous distribution function, and by Theorem 1.51 the set function given by

$$\mu(a,b] = F(b) - F(a) = \int_0^b f - \int_0^a f = \int_a^b f$$

for $-\infty < a < b < \infty$ has a unique extension to a Lebesgue-Stieltjes measure $\mu^* : \mathcal{B}(\mathbb{R}) \to [0, \infty]$.

The case when $f \equiv 1$ and $F(0) := 0$ is special. We then find that $F(x) = x$ for all $x \in \mathbb{R}$, so that $\mu(a,b] = b - a$ for $-\infty < a < b < \infty$, which of course conforms to the usual notion of the "length" of a bounded interval. The unique extension of μ in this case is denoted by λ_1, or

[2] See also §2.4.

simply λ if no ambiguity results,[3] and we call $\lambda_1 : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ the **Lebesgue measure** on $\mathcal{B}(\mathbb{R})$. Applying (1.11) through (1.15) to λ_1, for which F is the identity function, we find that

$$\lambda_1(a,b) = \lambda_1[a,b) = \lambda_1[a,b] = b - a$$

for all $a, b \in \mathbb{R}$,

$$\lambda_1(\mathbb{R}) = \lambda_1(-\infty, x) = \lambda_1(x,\infty) = \lambda_1(-\infty, x] = \lambda_1[x,\infty) = \infty$$

for all $x \in \mathbb{R}$, and finally

$$\lambda_1\{x\} = 0$$

for all $x \in \mathbb{R}$.

The completion of $\mathcal{B}(\mathbb{R})$ relative to λ_1 we denote by $\overline{\mathcal{B}}(\mathbb{R})$ rather than $\mathcal{B}(\mathbb{R})_{\lambda_1}$, and the complete measure $\overline{\lambda}_1 : \overline{\mathcal{B}}(\mathbb{R}) \to [0,\infty]$ we will often denote again by λ_1 (or simply λ) and also call **Lebesgue measure**. We call $\overline{\mathcal{B}}(\mathbb{R})$ the class of **Lebesgue measurable sets in** $\mathbb{R}$. ■

An element of $\mathbb{R}^n$ will be regarded as an $n \times 1$ matrix (i.e. a column vector), so that

$$\mathbb{R}^n = \Big\{ [x_i]_{n\times 1} : \forall 1 \le i \le n\big(x_i \in \mathbb{R}\big) \Big\},$$

where

$$[x_i]_{n\times 1} := \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Usually $[x_i]_{n\times 1}$ will be written simply as $[x_i]$ or $\mathbf{x}$.

Definition 1.53. *Given $\mathbf{a} = [a_i]_{n\times 1}$ and $\mathbf{b} = [b_i]_{n\times 1}$ in $\mathbb{R}^n$, we define the **h-interval** in $\mathbb{R}^n$ by*

$$(\mathbf{a},\mathbf{b}] = \prod_{i=1}^n (a_i, b_i] = \Big\{ [x_i] \in \mathbb{R}^n : \forall 1 \le i \le n\big(x_i \in (a_i, b_i]\big) \Big\},$$

*where $(\mathbf{a},\mathbf{b}] = \varnothing$ if $a_i \ge b_i$ for some i. Similarly we define **open intervals** and **closed intervals** in $\mathbb{R}^n$ by*

$$(\mathbf{a},\mathbf{b}) = \prod_{i=1}^n (a_i, b_i) \quad \text{and} \quad [\mathbf{a},\mathbf{b}] = \prod_{i=1}^n [a_i, b_i],$$

respectively, and also define

$$(-\infty,\mathbf{b}) = \prod_{i=1}^n (-\infty, b_i) \quad \text{and} \quad (\mathbf{a},\infty) = \prod_{i=1}^n (a_i, \infty).$$

Other intervals such as $(-\infty,\mathbf{b}]$ and $[\mathbf{a},\infty)$ are defined similarly.

Given $\mathbf{a},\mathbf{b} \in \mathbb{R}^n$, we write $\mathbf{a} < \mathbf{b}$ if $a_i < b_i$ for each $1 \le i \le n$, and we write $\mathbf{a} \le \mathbf{b}$ if $a_i \le b_i$ for each $1 \le i \le n$. Thus $(\mathbf{a},\mathbf{b}] \ne \varnothing$ if and only if $\mathbf{a} < \mathbf{b}$.

Letting

$$\mathcal{H}_n = \big\{ (\mathbf{a},\mathbf{b}] : \mathbf{a},\mathbf{b} \in \mathbb{R}^n \big\} \quad \text{and} \quad \overline{\mathcal{H}}_n = \big\{ (\mathbf{a},\mathbf{b}] : \mathbf{a},\mathbf{b} \in \overline{\mathbb{R}}^n \big\}$$

---

[3] See Example 1.56 for the definition of λ_n for any $n \in \mathbb{N}$.

it can be shown that the σ-algebra of **Borel sets** of $\mathbb{R}^n$ is generated by $\mathcal{H}_n$: $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{H}_n)$. Similarly for the Borel sets of $\overline{\mathbb{R}}^n = [-\infty, \infty]^n$ we have $\mathcal{B}(\overline{\mathbb{R}}^n) = \sigma(\overline{\mathcal{H}}_n)$.

A **Lebesgue-Stieltjes measure** on $\mathbb{R}^n$ is a measure $\mu : \mathcal{B}(\mathbb{R}^n) \to [0, \infty]$ such that $\mu(I) < \infty$ for every bounded interval $I \subseteq \mathbb{R}^n$. The notion of a distribution function on $\mathbb{R}^n$ does not carry over as easily, however.

Let $\Phi : \mathbb{R}^n \to \mathbb{R}$ be any function. For any $1 \le i \le n$ and $a, b \in \mathbb{R}$ define $\triangle_{b,a}^{(i)}\Phi : \mathbb{R}^n \to \mathbb{R}$ by

$$\triangle_{b,a}^{(i)}\Phi\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = \Phi\left(\begin{bmatrix} x_1 \\ \vdots \\ x_{i-1} \\ b \\ x_{i+1} \\ \vdots \\ x_n \end{bmatrix}\right) - \Phi\left(\begin{bmatrix} x_1 \\ \vdots \\ x_{i-1} \\ a \\ x_{i+1} \\ \vdots \\ x_n \end{bmatrix}\right) := \Phi\begin{bmatrix} x_1 \\ \vdots \\ x_{i-1} \\ b \\ x_{i+1} \\ \vdots \\ x_n \end{bmatrix} - \Phi\begin{bmatrix} x_1 \\ \vdots \\ x_{i-1} \\ a \\ x_{i+1} \\ \vdots \\ x_n \end{bmatrix},$$

where the last expression is a notational convenience. We take the **difference operator** $\triangle_{b,a}^{(i)}$ to have the additive linearity property

$$\triangle_{b,a}^{(i)}\big[\Phi(\mathbf{x}) \pm \Phi(\mathbf{y})\big] = \triangle_{b,a}^{(i)}\Phi(\mathbf{x}) \pm \triangle_{b,a}^{(i)}\Phi(\mathbf{y}).$$

Now, for any $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ the expression

$$\triangle_{b_1,a_1}^{(1)} \cdots \triangle_{b_n,a_n}^{(n)}\Phi(\mathbf{x})$$

is shown in [PMT] to be independent of the choice for $\mathbf{x}$, and so we define

$$\Phi(\mathbf{a}, \mathbf{b}] = \triangle_{b_1,a_1}^{(1)} \cdots \triangle_{b_n,a_n}^{(n)}\Phi(\mathbf{0}), \tag{1.16}$$

where $\mathbf{0}$ denotes the origin in $\mathbb{R}^n$. We say $\Phi : \mathbb{R}^n \to \mathbb{R}$ is **increasing** if $\Phi(\mathbf{a}, \mathbf{b}] \ge 0$ whenever $\mathbf{a} \le \mathbf{b}$.

Assume $\mathbb{R}^n$ to have the usual Euclidean metric. We say $\Phi : \mathbb{R}^n \to \mathbb{R}$ is **right-continuous** at $\mathbf{a}$ if

$$\lim_{\mathbf{x} \to \mathbf{a}^+} \Phi(\mathbf{x}) = \Phi(\mathbf{a}),$$

where $\mathbf{x} \to \mathbf{a}^+$ is taken to mean that $\mathbf{x} \to \mathbf{a}$ (i.e. $\mathbf{x}$ approaches $\mathbf{a}$) and $\mathbf{x} \ge \mathbf{a}$.

Finally, we define a function $F : \mathbb{R}^n \to \mathbb{R}$ to be a **distribution function** on $\mathbb{R}^n$ if it is increasing and right-continuous on $\mathbb{R}^n$. When $n = 1$ this definition agrees with the earlier definition of a distribution function on $\mathbb{R}$.

Proposition 1.54. *Let μ be a finite measure on $\mathcal{B}(\mathbb{R}^n)$ and define $\Phi : \mathbb{R}^n \to \mathbb{R}$ by*

$$\Phi(\mathbf{x}) = \mu(-\infty, \mathbf{x}].$$

If $\mathbf{a} \le \mathbf{b}$, then

$$\mu(\mathbf{a}, \mathbf{b}] = \Phi(\mathbf{a}, \mathbf{b}] = \sum_{k=0}^{n} (-1)^k s_k,$$

where s_k is the sum of all $\binom{n}{k}$ terms of the form $\Phi([c_i])$, the vector $[c_i]$ having k entries for which $c_i = a_i$, and $n - k$ entries for which $c_i = b_i$.

The following theorem becomes Theorem 1.51 when $n = 1$, and the general structure of the proof is much the same.

Theorem 1.55. *Let F be a distibution function on $\mathbb{R}^n$. If a set function μ is given by $\mu(\mathbf{a}, \mathbf{b}] = F(\mathbf{a}, \mathbf{b}]$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ with $\mathbf{a} \leq \mathbf{b}$, then μ has a unique extension to a Lebesgue-Stieltjes measure on $\mathbb{R}^n$.*

Example 1.56. If $F_1, \ldots, F_n$ are distribution functions on $\mathbb{R}$, then $F : \mathbb{R}^n \to \mathbb{R}$ given by

$$F(\mathbf{x}) = \prod_{i=1}^{n} F_i(x_i)$$

for all $\mathbf{x} = [x_i] \in \mathbb{R}^n$ is a distribution function on $\mathbb{R}^n$. Then for intervals $(\mathbf{a}, \mathbf{b}] \subseteq \mathbb{R}^n$ it can be shown that

$$F(\mathbf{a}, \mathbf{b}] = \Delta_{b_1,a_1}^{(1)} \cdots \Delta_{b_n,a_n}^{(n)} F(\mathbf{0}) = \prod_{i=1}^{n} \big[F_i(b_i) - F_i(a_i)\big].$$

Consider the $n = 2$ case:

$$\begin{aligned}
F(\mathbf{a}, \mathbf{b}] &= \Delta_{b_1,a_1}^{(1)} \Delta_{b_2,a_2}^{(2)} F\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \Delta_{b_1,a_1}^{(1)} \left(F\begin{bmatrix} 0 \\ b_2 \end{bmatrix} - F\begin{bmatrix} 0 \\ a_2 \end{bmatrix} \right) \\
&= \left(F\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} - F\begin{bmatrix} a_1 \\ b_2 \end{bmatrix} \right) - \left(F\begin{bmatrix} b_1 \\ a_2 \end{bmatrix} - F\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \right) \\
&= \big[F_1(b_1)F_2(b_2) - F_1(a_1)F_2(b_2)\big] - \big[F_1(b_1)F_2(a_2) - F_1(a_1)F_2(a_2)\big] \\
&= \big[F_1(b_1) - F_1(a_1)\big]F_2(b_2) - \big[F_1(b_1) - F_1(a_1)\big]F_2(a_2) \\
&= \big[F_1(b_1) - F_1(a_1)\big]\big[F_2(b_2) - F_2(a_2)\big]
\end{aligned}$$

If each $F_i : \mathbb{R} \to \mathbb{R}$ is the identity map, then $F(\mathbf{x}) = \prod_{i=1}^{n} x_i$ and we obtain

$$F(\mathbf{a}, \mathbf{b}] = \prod_{i=1}^{n} (b_i - a_i),$$

the customary volume of the interval $(\mathbf{a}, \mathbf{b}]$ which is, after all, a rectangular box in $\mathbb{R}^n$. The Lebesgue-Stieltjes measure μ on $\mathbb{R}^n$ that is induced (via Theorem 1.55) by the distribution function $F(\mathbf{x}) = \prod_{i=1}^{n} x_i$ is the **Lebesgue measure on** $\mathcal{B}(\mathbb{R}^n)$ and denoted by λ_n (or simply λ if no ambiguity results). In particular we have

$$\lambda_n(\mathbf{a}, \mathbf{b}] = \lambda_n\left(\prod_{i=1}^{n} (a_i, b_i]\right) = \prod_{i=1}^{n} (b_i - a_i) = \prod_{i=1}^{n} \lambda_1(a_i, b_i],$$

and it is routine to verify that

$$\lambda_n(\mathbf{a}, \mathbf{b}] = \lambda_n[\mathbf{a}, \mathbf{b}] = \lambda_n[\mathbf{a}, \mathbf{b}) = \lambda_n(\mathbf{a}, \mathbf{b}).$$

The completion of $\mathcal{B}(\mathbb{R}^n)$ relative to λ_n we will denote by $\overline{\mathcal{B}}(\mathbb{R}^n)$ rather than $\mathcal{B}(\mathbb{R}^n)_{\lambda_n}$, and the complete measure $\overline{\lambda_n} : \overline{\mathcal{B}}(\mathbb{R}^n) \to [0, \infty]$ we will often denote again by λ_n and also call **Lebesgue measure**. We call $\overline{\mathcal{B}}(\mathbb{R}^n)$ the class of **Lebesgue measurable sets on** $\mathbb{R}^n$. ∎

Remark. Some results that are true for complete Lebesgue measure do not apply to noncomplete Lebesgue measure, so there is risk associated with careless use of the symbol λ_n. If complete Lebesgue measure is an essential hypothesis in a proposition, then either the symbol $\overline{\lambda}_n$ will be employed, or else the measure space $(\mathbb{R}^n, \overline{\mathcal{B}}(\mathbb{R}^n), \lambda_n)$ will be declared. Put plainly: whenever the σ-algebra $\overline{\mathcal{B}}(\mathbb{R}^n)$ is under consideration, λ_n will be taken to denote complete Lebesgue measure.

Problem 1.57 (PMT 1.4.1). Let $F : \mathbb{R} \to \mathbb{R}$ be the distribution function given by

$$F(x) = \begin{cases} 0, & x \in (-\infty, -1) \\ 1 + x, & x \in [-1, 0) \\ 2 + x^2, & x \in [0, 2) \\ 9, & x \in [2, \infty), \end{cases}$$

and let μ be the Lebesgue-Stieltjes measure induced by F. Find the measure of each of following sets:

$$\{2\}, \quad \left[-\tfrac{1}{2}, 3\right), \quad (-1, 0] \cup (1, 2), \quad \{x : |x| + 2x^2 > 1\}.$$

Solution. By (1.15),

$$\mu(\{2\}) = F(2) - F(2^-) = 9 - (2 + 2^2) = 3,$$

and by (1.11),

$$\mu\left[-\tfrac{1}{2}, 3\right) = F(3^-) - F\left(-\tfrac{1}{2}^-\right) = 9 - \tfrac{1}{2} = \tfrac{17}{2}$$

and

$$\mu\big((-1, 0] \cup (1, 2)\big) = \mu(-1, 0] + \mu(1, 2) = \big[F(0) - F(-1)\big] + \big[F(2^-) - F(1)\big] = 2 + 3 = 5.$$

Finally, we have

$$\begin{aligned} |x| + 2x^2 = 1 \;&\Rightarrow\; x^2 = (2x^2 - 1)^2 \;\Rightarrow\; 4x^4 - 5x^2 + 1 = 0 \\ &\Rightarrow\; (4x^2 - 1)(x^2 - 1) = 0 \;\Rightarrow\; x = \pm\tfrac{1}{2}, \pm 1. \end{aligned}$$

Casting out extraneous solutions leaves $x = \pm\frac{1}{2}$. Now, the function $h(x) = |x| + 2x^2 - 1$ is continuous on $\mathbb{R}$, so to find where $h(x) > 0$ we determine the sign of $h(x)$ on the intevals $(-\infty, -\frac{1}{2})$, $(-\frac{1}{2}, \frac{1}{2})$, and $(\frac{1}{2}, \infty)$, and apply the Intermediate Value Theorem. We find that

$$\{x : |x| + 2x^2 > 1\} = \{x : h(x) > 0\} = \left(-\infty, -\tfrac{1}{2}\right) \cup \left(\tfrac{1}{2}, \infty\right)$$

Now, by (1.12) and (1.13),

$$\begin{aligned} \mu\big(\{x : |x| + 2x^2 > 1\}\big) &= \left[F\left(-\tfrac{1}{2}\right) - F(-\infty)\right] + \left[F(\infty) - F\left(\tfrac{1}{2}\right)\right] \\ &= \left(\tfrac{1}{2} - 0\right) + \left(9 - \tfrac{9}{4}\right) = \tfrac{29}{4} = 7\tfrac{1}{4}. \end{aligned}$$

■

Problem 1.58 (PMT 1.4.2). Let μ be a Lebesgue-Stieltjes measure on $\mathbb{R}$ induced by a continuous distribution function F.

(a) If A is a countable subset of $\mathbb{R}$, show that $\mu(A) = 0$.
(b) If $\mu(A) > 0$, must A include an open interval?

(c) If $\mu(A) > 0$ and $\mu(\mathbb{R} \setminus A) = 0$, must A be dense in $\mathbb{R}$?
(d) Do the answers to (b) or (c) change if μ is restricted to be Lebesgue measure?

Solution.
(a) Suppose $A \subseteq \mathbb{R}$ is countable, so $A = \{a_n : n \in \mathbb{N}\}$. First observe that $A \in \mathcal{B}(\mathbb{R})$ since $A = \bigcup_{n=1}^{\infty}\{a_n\}$, where $\{a_n\} \in \mathcal{B}(\mathbb{R})$ for each n and $\mathcal{B}(\mathbb{R})$ is a σ-algebra. Now, by countable additivity, (1.15), and the continuity of F,

$$\mu(A) = \mu\left(\bigcup_{n=1}^{\infty}\{a_n\}\right) = \sum_{n=1}^{\infty}\mu(\{a_n\}) = \sum_{n=1}^{\infty}\big[F(a_n) - F(a_n^-)\big] = \sum_{n=1}^{\infty}(0) = 0.$$

(b) If F is a constant function then $\mu(A) = 0$ for all $A \in \mathcal{B}(\mathbb{R})$, so suppose that F is not constant. Since $\mathbb{Q}$ is countable we have $\mathbb{Q} \in \mathcal{B}(\mathbb{R})$ with $\mu(\mathbb{Q}) = 0$. Now, $\mathbb{R} \setminus \mathbb{Q} \in \mathcal{B}(\mathbb{R})$ with

$$\mu(\mathbb{R} \setminus \mathbb{Q}) = \mu(\mathbb{R}) - \mu(\mathbb{Q}) = \mu(\mathbb{R}) = F(\infty) - F(-\infty) > 0$$

by Theorem 1.20(4) and equation (1.14). Recalling that $\mathbb{Q}$ is dense in $\mathbb{R}$, the set of irrational numbers $\mathbb{R} \setminus \mathbb{Q}$ contains no open interval, and so we conclude that $\mu(A) > 0$ does not necessarily imply that A contains an open inteval.

(c) Suppose $F : \mathbb{R} \to \mathbb{R}$ is given by

$$F(x) = \begin{cases} 0, & x \in (-\infty, 0] \\ x, & x \in (0, 1] \\ 1, & x \in (1, \infty), \end{cases}$$

which is of course a continuous function. Let $A = (0, 1]$. Then $\mu(A) = F(1) - F(0) = 1$, and

$$\begin{aligned} \mu(\mathbb{R} \setminus A) &= \mu(-\infty, 0] + \mu(1, \infty) = \big[F(0) - F(-\infty)\big] + \big[F(\infty) - F(1)\big] \\ &= (0 - 0) + (1 - 1) = 0. \end{aligned}$$

So $\mu(A) > 0$ while $\mu(\mathbb{R} \setminus A) = 0$, and yet A is not dense in $\mathbb{R}$.

(d) If λ is Lebesgue measure, so that $F(x) = x$ for all $x \in \mathbb{R}$ (i.e. F is not constant), then the conclusion of part (b) applies to λ.

The conclusion of part (c) does change, however. Suppose $A \in \mathcal{B}(\mathbb{R})$ is such that $\lambda(A) > 0$ and $\lambda(\mathbb{R} \setminus A) = 0$. Suppose A is not dense in $\mathbb{R}$. Then there exists some open interval $I \subseteq \mathbb{R}$ such that $A \cap I = \varnothing$. Thus $I \subseteq \mathbb{R} \setminus A$, and Theorem 1.20(5) implies that $\lambda(I) \leq \lambda(\mathbb{R} \setminus A) = 0$. Thus $\lambda(I) = 0$, which is impossible since I is an open interval. Therefore A is dense in $\mathbb{R}$. ■

Let V be a vector space. For any set $S \subseteq V$ and $v \in V$ we define sets

$$v + S = \{v + x : x \in S\} \quad \text{and} \quad -S = \{-x : x \in S\}.$$

Note in general that $x \in -S$ if and only if $-x \in S$.

Problem 1.59 (PMT 1.4.3).
(a) Show that $\mathbf{a} + B \in \mathcal{B}(\mathbb{R}^n)$ for all $\mathbf{a} \in \mathbb{R}^n$ and $B \in \mathcal{B}(\mathbb{R}^n)$.

(b) Show that $-B \in \mathcal{B}(\mathbb{R}^n)$ for all $B \in \mathcal{B}(\mathbb{R}^n)$.
(c) Show that $\lambda(\mathbf{a} + B) = \lambda(B)$ for all $B \in \mathcal{B}(\mathbb{R}^n)$.

Solution.
(a) Let $\mathbf{a} \in \mathbb{R}^n$, and define

$$\mathcal{S}_{\mathbf{a}} = \{B \in \mathcal{B}(\mathbb{R}^n) : \mathbf{a} + B \in \mathcal{B}(\mathbb{R}^n)\}.$$

Also let $\mathcal{H} = \{(\mathbf{x}, \mathbf{y}] : \mathbf{x}, \mathbf{y} \in \mathbb{R}^n\}$. For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ we have $(\mathbf{x}, \mathbf{y}] \in \mathcal{B}(\mathbb{R}^n)$, and it is easy to verify that

$$\mathbf{a} + (\mathbf{x}, \mathbf{y}] = (\mathbf{a} + \mathbf{x}, \mathbf{a} + \mathbf{y}] \in \mathcal{B}(\mathbb{R}^n),$$

and so $\mathcal{H} \subseteq \mathcal{S}_{\mathbf{a}}$. Also, since $\mathbf{a} + \mathbb{R}^n = \mathbb{R}^n$, it is clear that $\mathbb{R}^n \in \mathcal{S}_{\mathbf{a}}$.

Let $(B_n)_n \subseteq \mathcal{S}_{\mathbf{a}}$, so for all n we have $B_n \in \mathcal{B}(\mathbb{R}^n)$ and $\mathbf{a} + B_n \in \mathcal{B}(\mathbb{R}^n)$. It follows that $\bigcup_n B_n \in \mathcal{B}(\mathbb{R}^n)$ and $\bigcup_n (\mathbf{a} + B_n) \in \mathcal{B}(\mathbb{R}^n)$. Since

$$\bigcup_n (\mathbf{a} + B_n) = \mathbf{a} + \bigcup_n B_n,$$

it follows that $\mathbf{a} + \bigcup_n B_n \in \mathcal{B}(\mathbb{R}^n)$, and hence $\bigcup_n B_n \in \mathcal{S}_{\mathbf{a}}$.

Next, let $B \in \mathcal{S}_{\mathbf{a}}$, so $B \in \mathcal{B}(\mathbb{R}^n)$ and $\mathbf{a} + B \in \mathcal{B}(\mathbb{R}^n)$, and subsequently $B^c \in \mathcal{B}(\mathbb{R}^n)$ and $(\mathbf{a} + B)^c \in \mathcal{B}(\mathbb{R}^n)$. Now, since $(\mathbf{a} + B)^c = \mathbf{a} + B^c$, it follows that $\mathbf{a} + B^c \in \mathcal{B}(\mathbb{R}^n)$, and hence $B^c \in \mathcal{S}_{\mathbf{a}}$.

Thus $\mathcal{S}_{\mathbf{a}}$ is a σ-algebra such that $\mathcal{S}_{\mathbf{a}} \supseteq \mathcal{H}$, implying that $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{H}) \subseteq \mathcal{S}_{\mathbf{a}}$. Clearly $\mathcal{S}_{\mathbf{a}} \subseteq \mathcal{B}(\mathbb{R}^n)$, and so we find that $\mathcal{S}_{\mathbf{a}} = \mathcal{B}(\mathbb{R}^n)$. This shows that $\mathbf{a} + B$ is a Borel set for every Borel set B.

(b) Define

$$\mathcal{S} = \{B \in \mathcal{B}(\mathbb{R}^n) : -B \in \mathcal{B}(\mathbb{R}^n)\},$$

and let $\mathcal{H} = \{(\mathbf{a}, \mathbf{b}] : \mathbf{a}, \mathbf{b} \in \mathbb{R}^n\}$. Now, $\mathbf{x} \in -(\mathbf{a}, \mathbf{b}]$ implies $\mathbf{x} = -\mathbf{y}$ for some $\mathbf{y} = [y_i] \in (\mathbf{a}, \mathbf{b}]$, where

$$\begin{aligned} \mathbf{y} \in (\mathbf{a}, \mathbf{b}] \quad &\Leftrightarrow \quad \mathbf{y} \in \prod_{i=1}^{n} (a_i, b_i] \quad \Leftrightarrow \quad \forall i (a_i < y_i \le b_i) \quad \Leftrightarrow \quad \forall i \big(-y_i \in [-b_i, -a_i) \big) \\ &\Leftrightarrow \quad -\mathbf{y} \in \prod_{i=1}^{n} [-b_i, -a_i) = [-\mathbf{b}, -\mathbf{a}), \end{aligned}$$

so $\mathbf{x} \in [-\mathbf{b}, -\mathbf{a})$ and we see that $-(\mathbf{a}, \mathbf{b}] \subseteq [-\mathbf{b}, -\mathbf{a})$. The reverse containment is similarly verified, so that $-(\mathbf{a}, \mathbf{b}] = [-\mathbf{b}, -\mathbf{a})$ for any $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$. Thus if $(\mathbf{a}, \mathbf{b}] \in \mathcal{H}$ then $(\mathbf{a}, \mathbf{b}] \in \mathcal{B}(\mathbb{R}^n)$ and $-(\mathbf{a}, \mathbf{b}] = [-\mathbf{b}, -\mathbf{a}) \in \mathcal{B}(\mathbb{R}^n)$, so $(\mathbf{a}, \mathbf{b}] \in \mathcal{S}$ and we conclude that $\mathcal{H} \subseteq \mathcal{S}$. It remains to show that $\mathcal{S}$ is a σ-algebra. Since $-\mathbb{R}^n = \mathbb{R}^n$, it is clear that $\mathbb{R}^n \in \mathcal{S}$.

Let $(B_n)_n \subseteq \mathcal{S}$, so $B_n, -B_n \in \mathcal{B}(\mathbb{R}^n)$ for each n, and thus $\bigcup_n B_n, \bigcup_n (-B_n) \in \mathcal{B}(\mathbb{R}^n)$. Now,

$$\begin{aligned} x \in \bigcup_n (-B_n) \quad &\Leftrightarrow \quad \exists k \big(x \in -B_k \big) \quad \Leftrightarrow \quad \exists k \big(-x \in B_k \big) \\ &\Leftrightarrow \quad -x \in \bigcup_n B_n \quad \Leftrightarrow \quad x \in -\bigcup_n B_n \end{aligned}$$

shows that $-\bigcup_n B_n = \bigcup_n (-B_n)$. From $\bigcup_n B_n, -\bigcup_n B_n \in \mathcal{B}(\mathbb{R}^n)$ it follows that $\bigcup_n B_n \in \mathcal{S}$, and thus $\mathcal{S}$ is closed under countable unionization.

Let $B \in \mathcal{S}$, so $B, -B \in \mathcal{B}(\mathbb{R}^n)$, and hence $B^c, (-B)^c \in \mathcal{B}(\mathbb{R}^n)$. Now,

$$x \in (-B)^c \ \Leftrightarrow \ x \notin -B \ \Leftrightarrow \ -x \notin B \ \Leftrightarrow \ -x \in B^c \ \Leftrightarrow \ x \in -B^c$$

shows that $(-B)^c = -B^c$. From $B^c, -B^c \in \mathcal{B}(\mathbb{R}^n)$ it follows that $B^c \in SS$, and thus $\mathcal{S}$ is closed under complementation.

Thus $\mathcal{S}$ is a σ-algebra such that $\mathcal{S} \supseteq \mathcal{H}$, implying that $\mathcal{B}(\mathbb{R}^n) \subseteq \mathcal{S}$, and hence $\mathcal{S} = \mathcal{B}(\mathbb{R}^n)$. Therefore $-B$ is a Borel set for every Borel set B.

(c) Fix $\mathbf{a} = [a_i] \in \mathbb{R}^n$. For $m \in \mathbb{N}$ define $C_m \subseteq \mathbb{R}^n$ by $C_m = \prod_{i=1}^n (-m, m]$, and again let $\mathcal{H} = \{(\mathbf{x}, \mathbf{y}] : \mathbf{x}, \mathbf{y} \in \mathbb{R}^n\}$. By Proposition 1.15, $\mathcal{B}_m := \mathcal{B}(\mathbb{R}^n) \cap C_m$ is a σ-algebra over C_m with

$$\mathcal{B}_m = \mathcal{B}(\mathbb{R}^n) \cap C_m = \sigma(\mathcal{H}) \cap C_m = \sigma_{C_m}(\mathcal{H} \cap C_m).$$

(Since C_m is a Borel set we see that $\mathcal{B}_m$ is precisely the collection of Borel sets of $\mathbb{R}^n$ that are subsets of C_m.) Define

$$\mathcal{S}_m = \big\{B \in \mathcal{B}_m : \lambda(\mathbf{a} + B) = \lambda(B)\big\},$$

so $\mathcal{S}_m \subseteq \mathcal{B}_m \subseteq \mathcal{B}(\mathbb{R}^n)$. Let $\mathcal{A}_m \subseteq \mathcal{P}(C_m)$ be the algebra of finite disjoint unions of h-intervals in C_m. We will show that $\mathcal{S}_m$ is a monotone class and $\mathcal{S}_m \supseteq \mathcal{A}_m$. Once this is done, the Monotone Class Theorem implies that $\mathcal{S}_m \supseteq \sigma_{C_m}(\mathcal{A}_m)$, and then since $\mathcal{H} \cap C_m \subseteq \mathcal{A}_m$ it follows that $\mathcal{S}_m \supseteq \sigma_{C_m}(\mathcal{H} \cap C_m) = \mathcal{B}_m$. That is, $\mathcal{S}_m = \mathcal{B}_m$, which is to say every Borel set B of $\mathbb{R}^n$ that is a subset of C_m is such that $\lambda(\mathbf{a} + B) = \lambda(B)$, finishing the proof.

Note that $(C_m, \mathcal{B}_m, \lambda)$ is a finite measure space. Let $(A_k) \subseteq \mathcal{S}_m$ such that $A_k \uparrow A$. Then $\mathbf{a} + A_k \uparrow \mathbf{a} + A$, and since $\mathbf{a} + A_k \in \mathcal{B}(\mathbb{R}^n)$ for each k by part (a), we have $\mathbf{a} + A \in \mathcal{B}(\mathbb{R}^n)$ as well. Now,

$$\lim_{k\to\infty} \lambda(A_k) = \lambda(A) \quad \text{and} \quad \lim_{k\to\infty} \lambda(\mathbf{a} + A_k) = \lambda(\mathbf{a} + A)$$

by Theorem 1.21(1), and since $\lambda(\mathbf{a} + A_k) = \lambda(A_k)$ for all k, we obtain

$$\lambda(A) = \lim_{k\to\infty} \lambda(A_k) = \lim_{k\to\infty} \lambda(\mathbf{a} + A_k) = \lambda(\mathbf{a} + A),$$

and hence $A \in \mathcal{S}_m$. If $(A_k) \subseteq \mathcal{S}_m$ is such that $A_k \downarrow A$, then a similar argument using Theorem 1.21(2) will show that $A \in \mathcal{S}_m$, where $\lambda(A_1) \in \mathbb{R}$ holds since $A_1 \subseteq C_m$. Therefore $\mathcal{S}_m$ is a montone class.

Now, if $A \in \mathcal{A}_m$, so that $A = \bigsqcup_{j=1}^k (\mathbf{p}_j, \mathbf{q}_j] \subseteq C_m$, then clearly $A \in \mathcal{B}_m$. Moreover,

$$\mathbf{a} + A = \bigsqcup_{j=1}^k (\mathbf{a} + \mathbf{p}_j, \mathbf{a} + \mathbf{q}_j] \in \mathcal{B}(\mathbb{R}^n)$$

with

$$\lambda(\mathbf{a} + A) = \sum_{j=1}^k \lambda(\mathbf{a} + \mathbf{p}_j, \mathbf{a} + \mathbf{q}_j] = \sum_{j=1}^k \lambda(\mathbf{p}_j, \mathbf{q}_j] = \lambda(A),$$

and hence $A \in \mathcal{S}_m$. Therefore $\mathcal{A}_m \subseteq \mathcal{S}_m$. ∎

Problem 1.60 (PMT 1.4.4). Show that $\lambda(\mathbf{a} + B) = \lambda(B)$ for all $B \in \overline{\mathcal{B}}(\mathbb{R}^n)$ and $\mathbf{a} \in \mathbb{R}^n$.

Solution. Fix $\mathbf{a} \in \mathbb{R}^n$. Let $B \in \overline{\mathcal{B}}(\mathbb{R}^n)$, so that $B = A \cup N$ for some $B \in \mathcal{B}(\mathbb{R}^n)$ and $N \in \mathcal{P}(\mathcal{N}_\lambda)$. Now, there is some $M \in \mathcal{B}(\mathbb{R}^n)$ with $\lambda(M) = 0$ such that $N \subseteq M$, and $\lambda(\mathbf{a} + M) = \lambda(M)$ by Problem 1.59(c). Since $\mathbf{a}+N \subseteq \mathbf{a}+M$ and $\lambda(\mathbf{a}+M) = 0$, it follows that $\mathbf{a}+N \in \mathcal{P}(\mathcal{N}_\lambda) \subseteq \overline{\mathcal{B}}(\mathbb{R}^n)$ and so $\lambda(\mathbf{a} + N)$ is defined in $[0, \infty]$. Indeed,

$$0 \leq \lambda(\mathbf{a} + N) \leq \lambda(\mathbf{a} + M) = 0$$

shows that $\lambda(\mathbf{a} + N) = 0$. Now, since

$$\mathbf{a} + B = (\mathbf{a} + A) \cup (\mathbf{a} + N),$$

we again apply Problem 1.59(c) to obtain

$$\lambda(A) = \lambda(\mathbf{a} + A) \leq \lambda(\mathbf{a} + B) \leq \lambda(\mathbf{a} + A) + \lambda(\mathbf{a} + N) = \lambda(\mathbf{a} + A) = \lambda(A),$$

recalling Theorem 1.20(5) and 1.20(6). Therefore

$$\lambda(\mathbf{a} + B) = \lambda(A) = \lambda(B),$$

the last equality being merely an outcome of the definition of a complete measure as given in Theorem 1.39. ■

2

Integration

2.1 – Measurable Functions

Definition 2.1. *Let* $(X, \mathcal{M})$ *and* $(Y, \mathcal{N})$ *be measurable spaces. A function* $f : X \to Y$ *is* $(\mathcal{M}, \mathcal{N})$***-measurable*** *if* $f^{-1}(A) \in \mathcal{M}$ *for all* $A \in \mathcal{N}$*, in which case we write*

$$f : (X, \mathcal{M}) \to (Y, \mathcal{N}).$$

In particular, given a measurable space $(X, \mathcal{M})$*, a function* $f : X \to \mathbb{R}^n$ *is* ***Borel measurable*** *if* f *is* $(\mathcal{M}, \mathcal{B}(\mathbb{R}^n))$*-measurable.*[4] *The same applies if* $\mathbb{R}^n$ *is replaced by* $\overline{\mathbb{R}}^n$*.*

A function $f : \mathbb{R}^m \to \mathbb{R}^n$ is **Lebesgue measurable** if $f^{-1}(A) \in \overline{\mathcal{B}}(\mathbb{R}^m)$ for all $A \in \mathcal{B}(\mathbb{R}^n)$. That is, a function $\mathbb{R}^m \to \mathbb{R}^n$ is Lebesgue measurable if the preimage of every Borel set of $\mathbb{R}^n$ is a Lebesgue measurable set of $\mathbb{R}^m$. Clearly a Borel measurable function $\mathbb{R}^m \to \mathbb{R}^n$ is necessarily Lebesgue measurable, but Lebesgue measurability does not imply Borel measurability.

We come to our next bit of terminology. In the case when $A \in \mathcal{B}(\mathbb{R}^m)$, to say a function $f : A \to \mathbb{R}^n$ is **Borel measurable** means specifically that $f : (A, \mathcal{B}(A)) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. The same applies if every $\mathbb{R}$ is replaced by $\overline{\mathbb{R}}$.

Proposition 2.2. *Let* $(X, \mathcal{M})$ *be a measurable space. If* $f : X \to \mathbb{R}$ *is* $(\mathcal{M}, \mathcal{B}(\mathbb{R}))$*-measurable, then it is also* $(\mathcal{M}, \mathcal{B}(\overline{\mathbb{R}}))$*-measurable.*

Proof. Suppose $f : X \to \mathbb{R}$ is $(\mathcal{M}, \mathcal{B}(\mathbb{R}))$-measurable. Let $B \in \mathcal{B}(\overline{\mathbb{R}})$. Since $\mathcal{B}(\mathbb{R}) = \mathcal{B}(\overline{\mathbb{R}}) \cap \mathbb{R}$ by Example 1.17, we have $B \cap \mathbb{R} \in \mathcal{B}(\mathbb{R})$, and thus $f^{-1}(B \cap \mathbb{R}) \in \mathcal{M}$. Now, $f^{-1}(\mathbb{R}) = X$ since f is real-valued, and so

$$f^{-1}(B \cap \mathbb{R}) = f^{-1}(B) \cap f^{-1}(\mathbb{R}) = f^{-1}(B) \cap X = f^{-1}(B).$$

Therefore $f^{-1}(B) \in \mathcal{M}$ and we conclude that f is $(\mathcal{M}, \mathcal{B}(\overline{\mathbb{R}}))$-measurable. ■

Proposition 2.3. *Let* $(X, \mathcal{M})$ *and* $(Y, \mathcal{N})$ *be measurable spaces, and suppose* $\mathcal{N} = \sigma(\mathcal{C})$*. If* $f : X \to Y$ *is such that* $f^{-1}(A) \in \mathcal{M}$ *for all* $A \in \mathcal{C}$*, then* f *is* $(\mathcal{M}, \mathcal{N})$*-measurable.*

[4] More explicitly we may say f is **Borel measurable on** $(X, \mathcal{M})$.

Proof. Let

$$\mathcal{G} = \{A \in \mathcal{N} : f^{-1}(A) \in \mathcal{M}\}.$$

Clearly $\mathcal{G} \supseteq \mathcal{C}$, and also $Y \in \mathcal{G}$ since $f^{-1}(Y) = X \in \mathcal{M}$. If $(A_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{G}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{N}$ and

$$f^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right) = \bigcup_{n=1}^{\infty} f^{-1}(A_n) \in \mathcal{M}$$

since $f^{-1}(A_n) \in \mathcal{M}$ for each n, and hence $\bigcup_{n=1}^{\infty} A_n \in \mathcal{G}$. Finally, if $A \in \mathcal{G}$, then $A^c \in \mathcal{N}$ and

$$f^{-1}(A^c) = \left[f^{-1}(A)\right]^c \in \mathcal{M}$$

since $f^{-1}(A) \in \mathcal{M}$, and so $A^c \in \mathcal{G}$. Thus $\mathcal{G}$ is a σ-algebra containing $\mathcal{C}$, implying $\mathcal{G} \supseteq \sigma(\mathcal{C}) = \mathcal{N}$, and therefore $\mathcal{G} = \mathcal{N}$. We conclude that $f^{-1}(A) \in \mathcal{M}$ for all $A \in \mathcal{N}$, so f is $(\mathcal{M}, \mathcal{N})$-measurable. ■

Proposition 2.4. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. If $f : X \to Y$ is continuous, then f is $(\mathcal{B}(X), \mathcal{B}(Y))$-measurable.*

Proof. Suppose $f : X \to Y$ is continuous. Let $V \in \mathcal{T}_Y$, which is to say V is an open set in the space Y. Then $f^{-1}(V)$ is open in X by the definition of continuity, or equivalently $f^{-1}(V) \in \mathcal{T}_X$. Since $\mathcal{B}(X) = \sigma(\mathcal{T}_X)$, it follows that $f^{-1}(V) \in \mathcal{B}(X)$. So $f^{-1}(V) \in \mathcal{B}(X)$ for all $V \in \mathcal{T}_Y$, and since $\mathcal{B}(Y) = \sigma(\mathcal{T}_Y)$, Proposition 2.3 implies that f is $(\mathcal{B}(X), \mathcal{B}(Y))$-measurable. ■

It is natural to think of a constant function $f : X \to \{c\} \subseteq Y$ as being continuous, but if either X or Y has not been given a topological structure there is no possibility (in the present context) of meaningfully discussing continuity. Nevertheless we have the following result.

Proposition 2.5. *Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces. If $f : X \to Y$ is a constant function, then f is $(\mathcal{M}, \mathcal{N})$-measurable.*

Proof. Suppose $f : X \to Y$ is a constant function, so there exists some $c \in Y$ such that $f(x) = c$ for all $x \in X$. Let $A \in \mathcal{N}$. Then

$$f^{-1}(A) = \{x \in X : f(x) \in A\} = \begin{cases} X, & c \in A \\ \varnothing, & c \notin A. \end{cases}$$

Since $X, \varnothing \in \mathcal{M}$, it follows that f is $(\mathcal{M}, \mathcal{N})$-measurable. ■

Proposition 2.6. *Suppose $(X, \mathcal{M})$ is a measurable space and $f : X \to \mathbb{R}$. The following are equivalent.*

1. *f is Borel 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. *$f^{-1}((-\infty, a)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.*
5. *$f^{-1}((-\infty, a]) \in \mathcal{M}$ for all $a \in \mathbb{R}$.*

Proof.
(1) → (2). Suppose f is Borel measurable. Since $\mathcal{B}(\mathbb{R})$ contains all the open sets in $\mathbb{R}$, we have $f^{-1}(U) \in \mathcal{M}$ for every open $U \subseteq \mathbb{R}$, and hence $f^{-1}((a,\infty)) \in \mathcal{M}$ for every $a \in \mathbb{R}$.

(2) → (3). Suppose $f^{-1}((a,\infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$. Fix $a \in \mathbb{R}$. For all $n \in \mathbb{N}$ we have $f^{-1}(a - 1/n, \infty)) \in \mathcal{M}$, and thus by Proposition 1.7 it follows that

$$f^{-1}([a,\infty)) = f^{-1}\left(\bigcap_{n=1}^{\infty}(a-1/n,\infty)\right) = \bigcap_{n=1}^{\infty} f^{-1}((a-1/n,\infty)) \in \mathcal{M}.$$

(3) → (4). Suppose $f^{-1}([a,\infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$. Fix $a \in \mathbb{R}$. From $f^{-1}([a,\infty)) \in \mathcal{M}$ we have

$$f^{-1}((-\infty,a)) = f^{-1}([a,\infty)^c) = \left[f^{-1}([a,\infty))\right]^c \in \mathcal{M}.$$

(4) → (5). Suppose $f^{-1}((-\infty,a)) \in \mathcal{M}$ for all $a \in \mathbb{R}$. Fix $a \in \mathbb{R}$. For all $n \in \mathbb{N}$ we have $f^{-1}(-\infty, a+1/n)) \in \mathcal{M}$, and thus by Proposition 1.7 it follows that

$$f^{-1}((-\infty,a]) = f^{-1}\left(\bigcap_{n=1}^{\infty}(-\infty,a+1/n)\right) = \bigcap_{n=1}^{\infty} f^{-1}((-\infty,a+1/n)) \in \mathcal{M}.$$

(5) → (1). Suppose $f^{-1}((-\infty,a]) \in \mathcal{M}$ for all $a \in \mathbb{R}$. The collection $\mathcal{C} = \{(-\infty,a] : a \in \mathbb{R}\}$ is such that $\sigma(\mathcal{C}) = \mathcal{B}(\mathbb{R})$, and so it follows by Proposition 2.3 that f is Borel measurable. ∎

Proposition 2.7. *Let $(X,\mathcal{M})$ be a measurable space, let $(f_n : X \to \overline{\mathbb{R}})_{n\in\mathbb{N}}$ be a sequence of Borel measurable functions, and suppose $\lim_{n\to\infty} f_n(x)$ exists in $\overline{\mathbb{R}}$ for each $x \in X$. Then the function $f : X \to \overline{\mathbb{R}}$ given by*

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

is Borel measurable.

Proof. In Example 1.13 it was shown that $\overline{\mathcal{O}} = \{(a,\infty] : a \in \mathbb{R}\}$ generates $\mathcal{B}(\overline{\mathbb{R}})$, and so by Proposition 2.3 it is enough to show that $f^{-1}((a,\infty]) \in \mathcal{M}$ for all $a \in \mathbb{R}$. By Proposition 1.3, for arbitrary $a \in \mathbb{R}$,

$$\begin{aligned}
f^{-1}((a,\infty]) &= \{x \in X : f(x) \in (a,\infty]\} = \left\{x : \lim_{n\to\infty} f_n(x) > a\right\} \\
&= \left\{x : \exists \ell \in \mathbb{N}\ \exists m \in \mathbb{N}\ \forall n \ge m \left(f_n(x) > a + \ell^{-1}\right)\right\} \\
&= \bigcup_{\ell=1}^{\infty}\left\{x : \exists m \in \mathbb{N}\ \forall n \ge m \left(f_n(x) > a + \ell^{-1}\right)\right\} \\
&= \bigcup_{\ell=1}^{\infty}\left\{x : x \in f_n^{-1}((a+\ell^{-1},\infty]) \text{ for all but finitely many } n\right\} \\
&= \bigcup_{\ell=1}^{\infty} \liminf_n f_n^{-1}((a+\ell^{-1},\infty]) = \bigcup_{\ell=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} f_k^{-1}((a+\ell^{-1},\infty]),
\end{aligned}$$

and since $f_k^{-1}((a+\ell^{-1},\infty]) \in \mathcal{M}$ for each $\ell, k \in \mathbb{N}$, it follows that $f^{-1}((a,\infty]) \in \mathcal{M}$ and we're done. ∎

Proposition 2.8. *Let* $(X, \mathcal{M})$ *be a measurable space. If* $f_1, \ldots, f_n : X \to \overline{\mathbb{R}}$ *are Borel measurable functions, then the functions*

$$\bigvee_{k=1}^{n} f_k \quad \text{and} \quad \bigwedge_{k=1}^{n} f_k$$

are Borel measurable.

Proof. The statement of the proposition is clearly true when $n = 1$. Suppose $f_1, f_2 : X \to \overline{\mathbb{R}}$ are Borel measurable, and let $a \in \mathbb{R}$. Then

$$\begin{aligned}
(f_1 \vee f_2)^{-1}((a, \infty]) &= \{x : f_1(x) \vee f_2(x) > a\} \\
&= \{x : f_1(x) > a \text{ or } f_2(x) > a\} \\
&= \{x : f_1(x) > a\} \cup \{x : f_2(x) > a\} \\
&= f_1^{-1}((a, \infty]) \cup f_2^{-1}((a, \infty]) \in \mathcal{M}
\end{aligned}$$

since $f_1^{-1}((a, \infty]) \in \mathcal{M}$ and $f_2^{-1}((a, \infty]) \in \mathcal{M}$. Example 1.13 showed $\overline{\mathcal{O}} = \{(a, \infty] : a \in \mathbb{R}\}$ generates $\mathcal{B}(\overline{\mathbb{R}})$, and so Proposition 2.3 implies that $f_1 \vee f_2$ is Borel measurable. Similarly,

$$\begin{aligned}
(f_1 \wedge f_2)^{-1}((a, \infty]) &= \{x : f_1(x) \wedge f_2(x) > a\} \\
&= \{x : f_1(x) > a \text{ and } f_2(x) > a\} \\
&= \{x : f_1(x) > a\} \cap \{x : f_2(x) > a\} \\
&= f_1^{-1}((a, \infty]) \cap f_2^{-1}((a, \infty]) \in \mathcal{M}
\end{aligned}$$

implies $f_1 \wedge f_2$ is Borel measurable. Thus the statement of the proposition is true when $n = 2$.

Suppose the statement of the proposition is true for some arbitrary $n \in \mathbb{N}$. Suppose $(f_k : X \to \overline{\mathbb{R}})_{k=1}^{n+1}$ are Borel measurable. Let $g = \bigvee_{k=1}^{n} f_k$ and $h = \bigwedge_{k=1}^{n} f_k$. Then g and h are Borel measurable by hypothesis, and since the proposition holds when $n = 2$ we find that

$$\bigvee_{k=1}^{n+1} f_k = g \vee f_{n+1} \quad \text{and} \quad \bigwedge_{k=1}^{n+1} f_k = h \wedge f_{n+1}$$

are Borel measurable. ∎

Proposition 2.9. *Let* $(X, \mathcal{M})$ *be a measurable space, and let* $(f_n : X \to \overline{\mathbb{R}})_{n \in \mathbb{N}}$ *be a sequence of Borel measurable functions. Then the functions*

$$\sup\nolimits_n f_n, \quad \inf\nolimits_n f_n, \quad \limsup\nolimits_n f_n, \quad \liminf\nolimits_n f_n$$

are Borel measurable.

Proof. The sequence $(\bigvee_{k=1}^{n} f_k)_{n \in \mathbb{N}}$ is a sequence of Borel measurable functions $X \to \overline{\mathbb{R}}$ by Proposition 2.8, and so we may conclude by Proposition 2.7 that $g : X \to \overline{\mathbb{R}}$ given by

$$g(x) = \lim_{n \to \infty} \bigvee_{k=1}^{n} f_k(x)$$

is Borel measurable provided that the limit exists in $\overline{\mathbb{R}}$ for each $x \in X$. Indeed, by Proposition 1.1 we have

$$\lim_{n\to\infty} \bigvee_{k=1}^{n} f_k(x) = \sup_n f_n(x) \in \overline{\mathbb{R}}$$

for each $x \in X$, so g is Borel measurable, and moreover $g = \sup_n f_n$. The proof that $\inf_n f_n$ is Borel measurable is similar.

Next, for each $x \in X$ we have

$$\limsup_{n\to\infty} f_n(x) = \lim_{n\to\infty} \left(\sup_{k\geq n} f_k(x) \right) \in \overline{\mathbb{R}},$$

where

$$\left(\sup_{k\geq n} f_k \right)_{n\in\mathbb{N}}$$

is a sequence of Borel measurable functions $X \to \overline{\mathbb{R}}$, and therefore $\limsup_n f_n : X \to \overline{\mathbb{R}}$ is Borel measurable by Proposition 2.7. The proof that $\liminf_n f_n$ is Borel measurable is similar. ∎

Proposition 2.10. *If* $f : (X_1, \mathcal{M}_1) \to (X_2, \mathcal{M}_2)$ *and* $g : (X_2, \mathcal{M}_2) \to (X_3, \mathcal{M}_3)$, *then*

$$g \circ f : (X_1, \mathcal{M}_1) \to (X_3, \mathcal{M}_3).$$

Proof. Let $A \in \mathcal{M}_3$. Then $g^{-1}(A) \in \mathcal{M}_2$, and so $(g \circ f)^{-1}(B) = f^{-1}(g^{-1}(A)) \in \mathcal{M}_1$. ∎

Note that if $f : \mathbb{R}^\ell \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^n$ are *Lebesgue* measurable functions, then it does not follow that $g \circ f : \mathbb{R}^\ell \to \mathbb{R}^n$ is Lebesgue measurable.

Let $n \in \mathbb{N}$, and let $X_1, \ldots, X_n$ be nonempty sets. For each $1 \leq k \leq n$ the ***k*th projection map** $\Pi_k : \prod_{j=1}^n X_j \to X_k$ is given by

$$\Pi_k\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right) = x_k \quad \text{for each} \quad \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \prod_{j=1}^{n} X_j. \tag{2.1}$$

Letting $[\mathbf{x}]_k$ denote the kth component of $\mathbf{x} \in \prod_{j=1}^n X_j$, we may write simply $\Pi_k(\mathbf{x}) = [\mathbf{x}]_k$. Projection maps are continuous functions whenever each X_k represents a metric space (X_k, d_k) and $\prod_{k=1}^n X_k$ is given the usual product topology. This is the case when $X_k = \mathbb{R}$ for each k, where $\mathbb{R}$ is equipped with the Euclidean metric. Also each $\Pi_k : \overline{\mathbb{R}}^n \to \overline{\mathbb{R}}$ is continuous if $\overline{\mathbb{R}}$ has the topology defined in Example 1.13 and $\overline{\mathbb{R}}^n$ has the product topology that arises therefrom.

Theorem 2.11. *Let* $(X, \mathcal{M})$ *be a measurable space, let* $f : X \to \overline{\mathbb{R}}^n$, *and define* $f_i = \Pi_i \circ f$ *for each* $1 \leq i \leq n$. *Then* f *is Borel measurable if and only if* f_i *is Borel measurable for all* $1 \leq i \leq n$.

Proof. Suppose f is Borel measurable, and fix $1 \leq i \leq n$. Let $a_i, b_i \in \overline{\mathbb{R}}$ with $a_i < b_i$. We have

$$\Pi_i^{-1}\big([a_i, b_i]\big) = \big\{ \mathbf{x} \in \overline{\mathbb{R}}^n : x_i \in [a_i, b_i] \text{ and } \forall j \neq i (x_j \in \overline{\mathbb{R}}) \big\} = \prod_{j=1}^{n} [a_j, b_j],$$

where $[a_j, b_j] = [-\infty, \infty]$ for all $j \neq i$, and so $\Pi_i^{-1}([a_i, b_i]) \in \mathcal{B}(\overline{\mathbb{R}}^n)$. Since the collection

$$\overline{\mathcal{C}} = \left\{[a, b] : a, b \in \overline{\mathbb{R}}\right\}$$

generates $\mathcal{B}(\overline{\mathbb{R}})$ by Example 1.13, Proposition 2.3 implies that Π_i is Borel measurable, and therefore f_i is Borel measurable by Proposition 2.10.

For the proof of the converse, which is short, see [PMT]. ■

With Theorem 2.11 we easily obtain a generalization of Proposition 2.7 to functions with codomain $\overline{\mathbb{R}}^n$ for any $n \in \mathbb{N}$.

Theorem 2.12. *Let $(X, \mathcal{M})$ be a measurable space, let $(f_i : X \to \overline{\mathbb{R}}^n)_{i\in\mathbb{N}}$ be a sequence of Borel measurable functions, and suppose $\lim_{i\to\infty} f_i(x)$ exists in $\overline{\mathbb{R}}^n$ for each $x \in X$. Then the function $f : X \to \overline{\mathbb{R}}^n$ given by*

$$f(x) = \lim_{i\to\infty} f_i(x)$$

is Borel measurable.

Proof. For each $i \in \mathbb{N}$ and $1 \leq j \leq n$ define $f_{ij} : X \to \overline{\mathbb{R}}$ by $f_{ij} = \Pi_j \circ f_i$, which is Borel measurable by Theorem 2.11. Since $\lim_{i\to\infty} f_{ij}(x)$ exists in $\overline{\mathbb{R}}$ for each $x \in X$, Proposition 2.7 implies that the function $f_j : X \to \overline{\mathbb{R}}$ given by

$$g_j(x) = \lim_{i\to\infty} f_{ij}(x)$$

is Borel measurable. Now, by the continuity of the projection map Π_j we have

$$\lim_{i\to\infty} f_{ij}(x) = \lim_{i\to\infty} \Pi_j(f_i(x)) = \Pi_j\Big(\lim_{i\to\infty} f_i(x)\Big) = \Pi_j(f(x))$$

for all $x \in X$, so that $\Pi_j \circ f = g_j$, and hence $\Pi_j \circ f$ is Borel measurable for all $1 \leq j \leq n$. Therefore f is Borel measurable by Theorem 2.11. ■

Problem 2.13 (PMT 1.5.2). Suppose $f, g : (X, \mathcal{M}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, let $A \in \mathcal{M}$, and define $h : X \to \overline{\mathbb{R}}$ by

$$h(x) = \begin{cases} f(x), & x \in A \\ g(x), & x \in A^c. \end{cases}$$

Show that h is Borel measurable.

Solution. Let $\mathcal{O} = \{(a, b) : a, b \in \overline{\mathbb{R}}\}$, which generates $\mathcal{B}(\overline{\mathbb{R}})$. For any $(a, b) \in \mathcal{O}$,

$$\begin{aligned} h^{-1}((a, b)) &= \{x : h(x) \in (a, b)\} \\ &= \{x \in A : h(x) \in (a, b)\} \cup \{x \in A^c : h(x) \in (a, b)\} \\ &= \{x \in A : f(x) \in (a, b)\} \cup \{x \in A^c : g(x) \in (a, b)\} \\ &= \big[\{x : f(x) \in (a, b)\} \cap A\big] \cup \big[\{x : g(x) \in (a, b)\} \cap A^c\big] \\ &= \big[f^{-1}((a, b)) \cap A\big] \cup \big[g^{-1}((a, b)) \cap A^c\big], \end{aligned}$$

where $f^{-1}((a, b)), g^{-1}((a, b)) \in \mathcal{M}$ since $(a, b) \in \mathcal{B}(\overline{\mathbb{R}})$. Thus $h^{-1}((a, b)) \in \mathcal{M}$, and therefore h is Borel measurable by Proposition 2.3. ■

Problem 2.14 (PMT 1.5.4). Let $(X, \mathcal{M}, \mu)$ be a complete measure space, and let $A \in \mathcal{M}$ with $\mu(A) = 0$. If $f : (X, \mathcal{M}) \to (Y, \mathcal{N})$, $g : X \to Y$, and $g|_{A^c} = f|_{A^c}$, show that g is $(\mathcal{M}, \mathcal{N})$-measurable.

Solution. Suppose $f : (X, \mathcal{M}) \to (Y, \mathcal{N})$, $g : X \to Y$, and $g|_{A^c} = f|_{A^c}$. Fix $B \in \mathcal{N}$. Since $g^{-1}(B) \cap A \subseteq A$ and $(X, \mathcal{M}, \mu)$ is complete, we have $g^{-1}(B) \cap A \in \mathcal{M}$. Also, since $f(x) = g(x)$ for all $x \in A^c$,

$$g^{-1}(B) \cap A^c = \{x \in A^c : g(x) \in B\} = \{x \in A^c : f(x) \in B\} = f^{-1}(B) \cap A^c,$$

and so $g^{-1}(B) \cap A^c \in \mathcal{M}$ since f is $(\mathcal{M}, \mathcal{N})$-measurable. Now,

$$g^{-1}(B) = \big(g^{-1}(B) \cap A\big) \cup \big(g^{-1}(B) \cap A^c\big) \in \mathcal{M},$$

and therefore g is $(\mathcal{M}, \mathcal{N})$-measurable. ■

2.2 – Lebesgue Integral Definitions and Properties

Henceforth the following arithmetic conventions will be observed in $\overline{\mathbb{R}}$, where $+\infty$ may be denoted by ∞: for all $a \in \mathbb{R}$,

$$a \pm \infty = \pm\infty, \quad a/\pm\infty = 0, \quad \text{and} \quad a \cdot (\pm\infty) = \begin{cases} \pm\infty, & a > 0 \\ \mp\infty, & a < 0. \end{cases}$$

Also $\infty + \infty = \infty$, $-\infty - \infty = -\infty$, and finally

$$0 \cdot (\pm\infty) = 0. \tag{2.2}$$

All addition and multiplication operations are commutative. The following are undefined in $\overline{\mathbb{R}}$ and called **indeterminate forms**:

$$\infty - \infty, \quad -\infty + \infty, \quad \frac{\pm\infty}{\pm\infty}, \quad \frac{\mp\infty}{\pm\infty}, \quad \frac{\pm\infty}{0}, \quad \frac{0}{0}.$$

Of course, (2.2) is not usual in other areas of analysis, but it will result in a nicer theory for our purposes.

Let X be a set and let $A \in \mathcal{P}(X)$. The **indicator** of A is the function $\chi_A : X \to \mathbb{R}$ (or $\overline{\mathbb{R}}$) given by

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

Note that for any $B \subseteq \mathbb{R}$ (or $\overline{\mathbb{R}}$) we have

$$\chi^{-1}(B) = \begin{cases} X, & 0, 1 \in B \\ A, & 0 \notin B \text{ and } 1 \in B \\ A^c, & 0 \in B \text{ and } 1 \notin B \\ \varnothing, & 0, 1 \notin B. \end{cases}$$

This immediately implies the following.

Proposition 2.15. *If $(X, \mathcal{M})$ is a measurable space, then χ_A is Borel measurable on $(X, \mathcal{M})$ if and only if $A \in \mathcal{M}$.*

Definition 2.16. *Let $(X, \mathcal{M})$ be a measurable space. A function $\varphi : X \to \overline{\mathbb{R}}$ is **simple** if φ is Borel measurable on $(X, \mathcal{M})$ and $\varphi(X)$ is a finite set.*

Proposition 2.17. *Let $(X, \mathcal{M})$ be a measurable space, and let $\varphi : X \to \overline{\mathbb{R}}$. Then φ is simple if and only if φ is expressible as a finite linear combination of Borel measurable indicators.*

Proof. Let $\varphi : X \to \overline{\mathbb{R}}$ be a simple function, so that for some $m \in \mathbb{N}$ there exist distinct values $a_1, \ldots, a_m \in \overline{\mathbb{R}}$ such that $\varphi(X) = \{a_1, \ldots, a_m\}$. Since φ is Borel measurable and $\{a_i\} \in \mathcal{B}(\overline{\mathbb{R}})$ for each $1 \le i \le m$, we find that the sets $A_i = \varphi^{-1}(\{a_i\})$ are disjoint elements of $\mathcal{M}$ that form a partition of X. Clearly

$$\varphi = \sum_{i=1}^{m} a_i \chi_{A_i}, \tag{2.3}$$

where χ_{A_i} is a Borel measurable indicator by Proposition 2.15.

For the converse, suppose that $\varphi = \sum_{i=1}^{m} a_i\chi_{A_i}$ such that each indicator χ_{A_i} is Borel measurable. Then $A_1, \ldots, A_m \in \mathcal{M}$ by Proposition 2.15. Fix $a \in \mathbb{R}$. If $(a, \infty]$ contains none of the values $a_1, \ldots, a_m$, then $\varphi^{-1}((a,\infty]) = \varnothing \in \mathcal{M}$. Suppose

$$(a,\infty] \cap \{a_1, \ldots, a_m\} = \{a_{i_1}, \ldots, a_{i_r}\} \neq \varnothing.$$

Then

$$\varphi^{-1}\big((a,\infty]\big) = \varphi^{-1}\big(\{a_{i_1}, \ldots, a_{i_r}\}\big) = \bigcup_{k=1}^{r} \varphi^{-1}\big(\{a_{i_k}\}\big) = \bigcup_{k=1}^{r} A_{i_k} \in \mathcal{M},$$

and we see that $\varphi^{-1}(I) \in \mathcal{M}$ for every interval I in the collection $\overline{\mathcal{O}} = \{(a,\infty] : a \in \mathbb{R}\}$. In Example 1.13 it was established that $\sigma(\overline{\mathcal{O}}) = \mathcal{B}(\overline{\mathbb{R}}$, and so Proposition 2.3 implies that φ is Borel measurable. Therefore φ is a simple function. ■

Any finite linear combination of Borel measurable indicators that equals a simple function φ, such as that given in (2.3), is called a **representation** of φ. Other representations may be possible: for each $1 \leq i \leq m$ there may exist disjoint sets $A_{i1}, \ldots, A_{in_i} \in \mathcal{M}$ such that $\bigcup_{j=1}^{n_i} A_{ij} = A_i$, in which case

$$\varphi = \sum_{i=1}^{m} \sum_{j=1}^{n_i} a_i \chi_{A_{ij}}$$

is another representation. Thus there may be representations of the form

$$\varphi = \sum_{j=1}^{n} b_j \chi_{B_j}$$

in which the coefficients $b_1, \ldots, b_n \in \varphi(X)$ are not distinct. The **canonical representation** of a simple function $\varphi : X \to \overline{\mathbb{R}}$ is

$$\varphi = \sum_{k=1}^{p} c_k \chi_{C_k},$$

where $c_1, \ldots, c_p$ are the distinct elements of $\varphi(X) \setminus \{0\}$ (i.e. the *nonzero* values in the range of φ) and $C_k = \varphi^{-1}(\{c_k\})$.

Proposition 2.18. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $\varphi : X \to \overline{\mathbb{R}}$ be a simple function. If*

$$\sum_{i=1}^{m} a_i \chi_{A_i} \quad \text{and} \quad \sum_{j=1}^{n} b_j \chi_{B_j} \tag{2.4}$$

are representations of φ, then

$$\sum_{i=1}^{m} a_i \mu(A_i) = \sum_{j=1}^{n} b_j \mu(B_j), \tag{2.5}$$

provided the sums exist in $\overline{\mathbb{R}}$.

Proof. Suppose the sums in (2.4) are two representations for φ, and suppose the sums in (2.5) exist in $\overline{\mathbb{R}}$. Another representation for φ is

$$\sum_{i=1}^{m}\sum_{j=1}^{n} t_{ij}\chi_{A_i\cap B_j},$$

where $t_{ij} = a_i = b_j$. Noting that the sets $A_i \cap B_j$ are disjoint and $\bigcup_{i=1}^{m} A_i = \bigcup_{j=1}^{n} B_j$, we obtain

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n} t_{ij}\mu(A_i\cap B_j) &= \sum_{i=1}^{m} a_i \sum_{j=1}^{n}\mu(A_i\cap B_j) = \sum_{i=1}^{m} a_i\,\mu\left(\bigcup_{j=1}^{n}(A_i\cap B_j)\right)\\ &= \sum_{i=1}^{m} a_i\,\mu\left(A_i\cap\bigcup_{j=1}^{n} B_j\right) = \sum_{i=1}^{m} a_i\mu(A_i),\end{aligned}$$

and

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n} t_{ij}\mu(A_i\cap B_j) &= \sum_{j=1}^{n}\sum_{i=1}^{m} t_{ij}\mu(A_i\cap B_j) = \sum_{j=1}^{n} b_j\sum_{i=1}^{m}\mu(A_i\cap B_j)\\ &= \sum_{j=1}^{n} b_j\,\mu\left(\bigcup_{i=1}^{m}(A_i\cap B_j)\right) = \sum_{j=1}^{n} b_j\,\mu\left(B_j\cap\bigcup_{i=1}^{m} A_i\right) = \sum_{j=1}^{n} b_j\mu(B_j),\end{aligned}$$

and therefore

$$\sum_{i=1}^{m} a_i\mu(A_i) = \sum_{i=1}^{m}\sum_{j=1}^{n} t_{ij}\mu(A_i\cap B_j) = \sum_{j=1}^{n} b_j\mu(B_j).$$

■

Definition 2.19. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $\varphi : X \to \overline{\mathbb{R}}$ be a simple function with representation $\sum_{i=1}^{m} a_i\chi_{A_i}$. The* ***Lebesgue integral of*** φ ***with respect to*** μ *is*

$$\int_X \varphi\, d\mu = \sum_{i=1}^{m} a_i\mu(A_i), \tag{2.6}$$

provided that the sum exists in $\overline{\mathbb{R}}$.

Proposition 2.18 ensures that the Lebesgue integral of a simple function is well-defined; that is, the value of the Lebesgue integral of a simple function φ is independent of the choice of representation for φ. If the sum in (2.6) does not exist as an extended real number, then we say the Lebesgue integral does not exist. Note, however, that if the simple function φ is nonnegative then (2.6) must exist in $[0, \infty]$.

Definition 2.20. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f : X \to \overline{\mathbb{R}}$ be a nonnegative Borel measurable function. The* ***Lebesgue integral of*** f ***with respect to*** μ *is*

$$\int_X f\, d\mu = \sup\left\{\int_X \varphi\, d\mu : \varphi \text{ is simple and } 0 \le \varphi \le f\right\}.$$

Unlike the Lebesgue integral of a simple function, the Lebesgue integral of a nonnegative Borel measurable function always exists in $\overline{\mathbb{R}}$.

For $f : X \to \overline{\mathbb{R}}$ define $f^+, f^- : X \to [0, \infty]$ by $f^+ = f \vee 0$ and $f^- = (-f) \vee 0$. Thus

$$f^+(x) = \max\{f(x), 0\} \quad \text{and} \quad f^-(x) = \max\{-f(x), 0\}$$

for all $x \in X$. It is straightforward to check that

$$f = f^+ - f^- \quad \text{and} \quad |f| = f^+ + f^-.$$

We call f^+ the **positive part** of f, and f^- the **negative part** of f.

Proposition 2.21. *Let $(X, \mathcal{M})$ be a measurable space. If $f : X \to \overline{\mathbb{R}}$ is Borel measurable, then so are f^+ and f^-.*

Proof. Suppose $f : X \to \overline{\mathbb{R}}$ is Borel measurable. The zero function $0 : X \to \{0\}$ is Borel measurable by Proposition 2.5, and so f^+ and f^- are Borel measurable by Proposition 2.8. ■

We may now easily extend the definition of the Lebesgue integral to include nearly every kind of Borel measurable function.

Definition 2.22. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f : X \to \overline{\mathbb{R}}$ be a Borel measurable function. The **Lebesgue integral of f with respect to μ** is*

$$\int_X f\,d\mu = \int_X f^+\,d\mu - \int_X f^-\,d\mu, \tag{2.7}$$

*provided the expression at right exists in $\overline{\mathbb{R}}$. If the expression at right exists in $\mathbb{R}$, we say f is **μ-integrable**.*[5]

The only way the Lebesgue integral in Definition 2.22 can fail to exist is if the expression at right in (2.7) assumes the indeterminate form $\infty - \infty$.

Definition 2.23. *Let $(X, \mathcal{M}, \mu)$ be a measure space, let $A \in \mathcal{M}$, and let $f : X \to \overline{\mathbb{R}}$ be a Borel measurable function. The **Lebesgue integral of f on A** (with respect to μ) is*

$$\int_A f\,d\mu = \int_X f\chi_A\,d\mu.$$

Remark. Note that for any Borel measurable function $f : X \to \overline{\mathbb{R}}$ we have, by Definitions 2.23 and 2.19,

$$\int_\varnothing f\,d\mu = \int_X f\chi_\varnothing\,d\mu = \int_X 0\,d\mu = 0\mu(\varnothing) = 0.$$

In particular $f\chi_\varnothing = 0f \equiv 0$ must be the case owing to our definition $0 \cdot (\pm\infty) = 0$ given in (2.2).

Theorem 2.24. *Let $(X, \mathcal{M})$ be a measurable space.*

[5] More explicitly we may say f is **μ-integrable on** $(X, \mathcal{M}, \mu)$.

1. *If $f : X \to [0, \infty]$ is a nonnegative Borel measurable function, then there is a sequence of nonnegative finite simple functions $(\varphi_n : X \to [0, \infty))_{n\in\mathbb{N}}$ such that*
$$\forall n \in \mathbb{N}\big(0 \le \varphi_n \le \varphi_{n+1} \le f\big) \quad \textit{and} \quad \forall x \in X\Big(\lim_{n\to\infty} \varphi_n(x) = f(x)\Big).$$
Moreover, if f is bounded then $(\varphi_n)_{n\in\mathbb{N}}$ converges uniformly to f.
2. *If $f : X \to \overline{\mathbb{R}}$ is an arbitrary Borel measurable function, then there is a sequence of finite simple functions $(\varphi_n : X \to \mathbb{R})_{n\in\mathbb{N}}$ such that*
$$\forall n \in \mathbb{N}\big(|\varphi_n| \le |f|\big) \quad \textit{and} \quad \forall x \in X\Big(\lim_{n\to\infty} \varphi_n(x) = f(x)\Big).$$

Proof.
Proof of Part (1). Let $f : X \to [0, \infty]$ be a nonnegative Borel measurable function. Fix $n \in \mathbb{N}$, and define
$$I_k = \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)$$
for each $k \in \{1, 2, \ldots, n2^n\}$. Now define $\varphi_n : X \to [0, \infty)$ by
$$\varphi_n(x) = \begin{cases} 2^{-n}(k-1), & x \in f^{-1}(I_k) \text{ for } k = 1, \ldots, n2^n \\ n, & x \in f^{-1}\big([n, \infty]\big). \end{cases}$$
Setting $I_{n2^n+1} = [n, \infty]$, we express φ_n in terms of indicators:
$$\varphi_n = \sum_{k=1}^{n2^n+1} \frac{k-1}{2^n} \chi_{f^{-1}(I_k)}.$$
The Borel measurability of f ensures that $f^{-1}(I_k) \in \mathcal{M}$ for all k, so that each indicator $\chi_{f^{-1}(I_k)}$ is Borel measurable by Proposition 2.15, and hence φ_n is a simple function by Proposition 2.17.

Fix $x \in X$, and fix $n \in \mathbb{N}$. Suppose $f(x) \in I_k$ for some $k \in \{1, 2, \ldots, n2^n\}$, so that $\varphi_n(x) = 2^{-n}(k-1)$. Then
$$f(x) \in \left[\frac{2(k-1)}{2^{n+1}}, \frac{2k}{2^{n+1}}\right),$$
and so either
$$f(x) \in \left[\frac{2k-2}{2^{n+1}}, \frac{2k-1}{2^{n+1}}\right) \quad \text{or} \quad f(x) \in \left[\frac{2k-1}{2^{n+1}}, \frac{2k}{2^{n+1}}\right).$$
If the former is the case, then
$$f(x) \ge \varphi_{n+1}(x) = \frac{2k-2}{2^{n+1}} = \frac{k-1}{2^n} = \varphi_n(x);$$
and if the latter, then
$$f(x) \ge \varphi_{n+1}(x) = \frac{2k-1}{2^{n+1}} > \frac{2k-2}{2^{n+1}} = \frac{k-1}{2^n} = \varphi_n(x).$$
Hence
$$\varphi_n(x) \le \varphi_{n+1}(x) \le f(x) \tag{2.8}$$
if $x \in f^{-1}([0, n))$. Another possibility is that $x \in [n + 1, \infty]$, in which case
$$\varphi_n(x) = n < n + 1 = \varphi_{n+1}(x) \le f(x),$$

and so (2.8) again holds if $x \in f^{-1}([n+1,\infty])$. The remaining possibility is that $x \in [n, n+1)$, but it is easy to check that the same conclusion (2.8) holds once more. Hence $0 \leq \varphi_n \leq \varphi_{n+1} \leq f$ holds for all $n \in \mathbb{N}$.

Again fix $x \in X$. Let $\epsilon > 0$, and choose $n_0 \in \mathbb{N}$ such that $2^{-n_0} < \epsilon$ and $f(x) \in [0, n_0)$. Thus there exists some $1 \leq k \leq n_0 2^{n_0}$ such that

$$x \in f^{-1}\left(\left[\frac{k-1}{2^{n_0}}, \frac{k}{2^{n_0}}\right)\right),$$

so that

$$0 \leq \varphi_{n_0}(x) = \frac{k-1}{2^{n_0}} \leq f(x) < \frac{k}{2^{n_0}} \quad \Rightarrow \quad \left|\varphi_{n_0}(x) - f(x)\right| < \frac{1}{2^{n_0}} < \epsilon.$$

Since $\varphi_{n_0}(x) \leq \varphi_n(x) \leq f(x)$ for all $n \geq n_0$, it follows that

$$\left|\varphi_n(x) - f(x)\right| < \epsilon$$

for all $n \geq n_0$, and therefore $\lim_{n\to\infty} \varphi_n(x) = f(x)$ for each $x \in X$.

Finally, suppose that f is bounded, so there exists some $M \in [0,\infty)$ such that $f \leq M$. Again let $\epsilon > 0$, and this time choose $n_0 \in \mathbb{N}$ such that $2^{-n_0} < \epsilon$ and $n_0 > M$. Note that $f^{-1}([m,\infty]) = \varnothing$. It is easy to check that $|\varphi_n(x) - f(x)| < 2^{-n} < \epsilon$ for all $n \geq n_0$ and $x \in X$, and therefore $(\varphi_n)_{n\in\mathbb{N}}$ converges uniformly to f.

Proof of Part (2). The proof obtains easily through use of Part (1), and is done in [PMT]. ■

Theorem 2.25. *Let $(X, \mathcal{M})$ be a measurable space.*

1. *If $f, g : X \to \overline{\mathbb{R}}^n$ are Borel measurable, then so is $f+g$ provided $\Pi_i(f(x)) + \Pi_i(g(x))$ is not an indeterminant form for any $1 \leq i \leq n$ and $x \in X$.*
2. *If $f, g : X \to \overline{\mathbb{R}}^n$ are Borel measurable, then so is $f-g$ provided $\Pi_i(f(x)) - \Pi_i(g(x))$ is not an indeterminant form for any $1 \leq i \leq n$ and $x \in X$.*
3. *$f, g : X \to \overline{\mathbb{R}}$ are Borel measurable, then so is fg provided $f(x)g(x)$ is not an indeterminant form for any $x \in X$.*
4. *$f, g : X \to \overline{\mathbb{R}}$ are Borel measurable, then so is f/g provided $f(x)/g(x)$ is not an indeterminant form for any $x \in X$.*

Proposition 2.26. *Let $(X, \mathcal{M})$ be a measurable space. If $f : X \to \overline{\mathbb{R}}$ is Borel measurable, then so is $|f|$.*

Proof. Suppose $f : X \to \overline{\mathbb{R}}$ is Borel measurable. Proposition 2.21 implies that f^+ and f^- are Borel measurable, and since $|f| = f^+ + f^-$, which never takes an indeterminant form, it follows by Theorem 2.25(1) that $|f|$ is Borel measurable. ■

Theorem 2.27. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and suppose $f, g : X \to \overline{\mathbb{R}}$ are Borel measurable functions such that $\int_X f\,d\mu$ and $\int_X g\,d\mu$ exist in $\overline{\mathbb{R}}$. The following properties hold.*

1. *If $c \in \mathbb{R}$ then $\int_X cf\,d\mu$ exists, with*

$$\int_X cf\,d\mu = c\int_X f\,d\mu.$$

2. *If $f \leq g$ then*

$$\int_X f\,d\mu \leq \int_X g\,d\mu.$$

3. *The integral $\int_X |f|\,d\mu$ exists, with*

$$\left|\int_X f\,d\mu\right| \leq \int_X |f|\,d\mu.$$

4. *If $f \geq 0$ and $A \in \mathcal{M}$, then*

$$\int_A f\,d\mu = \sup\left\{\int_A \varphi\,d\mu : \varphi \text{ is simple and } 0 \leq \varphi \leq f\right\}.$$

5. *The integral $\int_A f\,d\mu$ exists for each $A \in \mathcal{M}$, and moreover $\int_A f\,d\mu$ is finite for each $A \in \mathcal{M}$ if $\int_X f\,d\mu$ is finite.*

Proposition 2.28. *If $(X, \mathcal{M}, \mu)$ is a finite measure space and $f : X \to \mathbb{R}$ is a bounded Borel measurable function, then $\int_X f\,d\mu$ exists in $\mathbb{R}$.*

Proof. Suppose $(X, \mathcal{M}, \mu)$ is a finite measure space and $f : X \to \mathbb{R}$ is a bounded Borel measurable function. Using Proposition 2.21, we find f^+ and f^- to be nonnegative bounded Borel measurable functions. Let $c \in \mathbb{R}$ be such that $f^+, f^- \leq c$ on X. Certainly $\int_X f^+\,d\mu$ and $\int_X f^-\,d\mu$ exist in $\overline{\mathbb{R}}$, and since by Theorem 2.27(2) and Definition 2.19 we have

$$\int_X f^+\,d\mu \leq \int_X c\,d\mu = c\mu(X) < \infty \quad \text{and} \quad \int_X f^-\,d\mu \leq \int_X c\,d\mu = c\mu(X) < \infty,$$

it must be that both integrals are real-valued. The desired conclusion now follows directly from Definition 2.22. ■

2.3 – Basic Integration Theorems

Theorem 2.29. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f : X \to \overline{\mathbb{R}}$ be a Borel measurable function such that $\int_X f\,d\mu$ exists. If $(S_n)_{n\in\mathbb{N}}$ is a disjoint sequence in $\mathcal{M}$, then*

$$\int_{\bigsqcup_{n=1}^{\infty} S_n} f\,d\mu = \sum_{n=1}^{\infty} \int_{S_n} f\,d\mu. \tag{2.9}$$

Proof. Let $S = \bigsqcup_{n=1}^{\infty} S_n$. First suppose that f is a nonnegative simple function. Then f is expressible as a finite linear combination of Borel measurable indicators by Proposition 2.17,

$$f = \sum_{i=1}^{m} a_i \chi_{A_i},$$

where $A_i \in \mathcal{M}$ and $a_i \in (0, \infty)$ for each $1 \le i \le m$, and

$$\int_X f\,d\mu = \sum_{i=1}^{m} a_i \mu(A_i)$$

must exist in $[0, \infty]$ since $f \ge 0$. Then by Definitions 2.23 and 2.19, and the countable additivity of μ,

$$\begin{aligned}
\int_{\bigsqcup_{n=1}^{\infty} S_n} f\,d\mu &= \int_X f\chi_S\,d\mu = \int_X \sum_{i=1}^{m} a_i\chi_{A_i}\chi_S\,d\mu = \int_X \sum_{i=1}^{m} a_i\chi_{A_i\cap S}\,d\mu \\
&= \sum_{i=1}^{m} a_i\mu(A_i\cap S) = \sum_{i=1}^{m}\left(a_i \sum_{n=1}^{\infty} \mu(A_i\cap S_n)\right) = \sum_{n=1}^{\infty}\sum_{i=1}^{m} a_i\mu(A_i\cap S_n) \\
&= \sum_{n=1}^{\infty}\left(\int_X \sum_{i=1}^{m} a_i\chi_{A_i\cap S_n}\,d\mu\right) = \sum_{n=1}^{\infty}\left(\int_X \sum_{i=1}^{m} a_i\chi_{A_i}\chi_{S_n}\,d\mu\right) \\
&= \sum_{n=1}^{\infty}\left(\int_X f\chi_{S_n}\,d\mu\right) = \sum_{n=1}^{\infty}\int_{S_n} f\,d\mu,
\end{aligned}$$

and so (2.9) holds if f is a nonnegative simple function.

Now suppose that f is a nonnegative Borel measurable function such that $\int_X f\,d\mu$ exists. Let φ be a simple function such that $0 \le \varphi \le f$. Then by Theorem 2.27(2),

$$\int_S \varphi\,d\mu = \sum_{n=1}^{\infty}\int_{S_n} \varphi\,d\mu \le \sum_{n=1}^{\infty}\int_{S_n} f\,d\mu,$$

and so

$$\int_S f\,d\mu = \sup\left\{\int_S \varphi\,d\mu : \varphi \text{ is simple and } 0 \le \varphi \le f\right\} \le \sum_{n=1}^{\infty}\int_{S_n} f\,d\mu$$

by Theorem 2.27(4).

If $\int_{S_k} f\,d\mu = \infty$ for some $k \in \mathbb{N}$, then since $\chi_S \ge \chi_{S_k}$ it follows by Theorem 2.27(2) that

$$\int_S f\,d\mu = \int_X f\chi_S\,d\mu \ge \int_X f\chi_{S_k}\,d\mu = \int_{S_k} f\,d\mu = \infty,$$

and so

$$\int_S f\,d\mu = \sum_{n=1}^{\infty} \int_{S_n} f\,d\mu = \infty$$

Thus we may assume that $\int_{S_n} f\,d\mu \in [0,\infty)$ for all n. Let $\epsilon > 0$. Fix $n \in \mathbb{N}$. Applying Theorem 2.27(4), for each $1 \le k \le n$ there exists a simple function $0 \le \varphi_k \le f$ such that

$$\int_{S_k} f\,d\mu - \frac{\epsilon}{n} \le \int_{S_k} \varphi_k\,d\mu \le \int_{S_k} f\,d\mu.$$

Let $\varphi = \bigvee_{k=1}^n \varphi_k$. Then φ is Borel measurable by Proposition 2.8, and since $\varphi(X) \subseteq \bigcup_{k=1}^n \varphi_k(X)$ it is clear that φ is a nonnegative simple function. Moreover, by Theorem 2.27(2) we have

$$\int_{S_k} f\,d\mu - \frac{\epsilon}{n} \le \int_{S_k} \varphi\,d\mu \le \int_{S_k} f\,d\mu$$

for all $1 \le k \le n$. Equation (2.9) applies to φ and finite disjoint sequences, so that

$$\begin{aligned} \int_{\bigsqcup_{k=1}^n S_k} f\,d\mu &\ge \int_{\bigsqcup_{k=1}^n S_k} \varphi\,d\mu = \sum_{k=1}^n \int_{S_k} \varphi\,d\mu \\ &\ge \sum_{k=1}^n \left(\int_{S_k} f\,d\mu - \frac{\epsilon}{n} \right) = \sum_{k=1}^n \int_{S_k} f\,d\mu - \epsilon, \end{aligned}$$

whereas

$$\int_S f\,d\mu = \int_X f\chi_S\,d\mu \ge \int_X f\chi_{\bigsqcup_{k=1}^n S_k}\,d\mu = \int_{\bigsqcup_{k=1}^n S_k} f\,d\mu.$$

Thus

$$\sum_{k=1}^n \int_{S_k} f\,d\mu - \epsilon \le \nu(S),$$

which implies

$$\sum_{k=1}^\infty \int_{S_k} f\,d\mu - \epsilon - \epsilon \le \nu(S)$$

since $n \in \mathbb{N}$ is arbitrary, and therefore

$$\sum_{k=1}^\infty \int_{S_k} f\,d\mu \le \int_S f\,d\mu$$

since $\epsilon > 0$ is arbitrary. Hence (2.9) holds if f is a nonnegative Borel measurable function.

Finally, let f be an arbitrary Borel measurable function such that $\int_X f\,d\mu$ exists. Then $\int_A f\,d\mu$ exists for all $A \in \mathcal{M}$ by Theorem 2.27(5), which by Definition 2.22 implies the existence of $\int_A f^+\,d\mu$ and $\int_A f^-\,d\mu$ in $\overline{\mathbb{R}}$ such that

$$\int_A f^+\,d\mu - \int_A f^-\,d\mu$$

is well-defined in $\overline{\mathbb{R}}$ as well (i.e. the two integrals in the difference cannot both equal ∞ simultaneously). Now, by Proposition 2.21 both f^+ and f^- are nonnegative Borel measurable functions, and so

$$\int_{\bigsqcup_{n=1}^{\infty} S_n} f\,d\mu = \int_{\bigsqcup_{n=1}^{\infty} S_n} f^+\,d\mu - \int_{\bigsqcup_{n=1}^{\infty} S_n} f^-\,d\mu = \sum_{n=1}^{\infty}\int_{S_n} f^+\,d\mu - \sum_{n=1}^{\infty}\int_{S_n} f^-\,d\mu$$
$$= \sum_{n=1}^{\infty}\left(\int_{S_n} f^+\,d\mu - \int_{S_n} f^-\,d\mu\right) = \sum_{n=1}^{\infty}\int_{S_n} f\,d\mu$$

as desired. ■

Corollary 2.30. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f : X \to \overline{\mathbb{R}}$ be a Borel measurable function such that $\int_X f\,d\mu$ exists. Define $\nu : \mathcal{M} \to \overline{\mathbb{R}}$ by*

$$\nu(A) = \int_A f\,d\mu$$

for all $A \in \mathcal{M}$. Then the following hold.

1. *The set function ν is countably additive on $\mathcal{M}$.*
2. *$(X, \mathcal{M}, \nu)$ is a measure space if $f \geq 0$.*
3. *If $\nu^+, \nu^- : \mathcal{M} \to \overline{\mathbb{R}}$ are given by*

$$\nu^+(A) = \int_A f^+\,d\mu \quad \text{and} \quad \nu^-(A) = \int_A f^-\,d\mu$$

for all $A \in \mathcal{M}$, then ν^+ and ν^- are measures such that $\nu = \nu^+ - \nu^-$.

Proof. Only the third statement requires comment. Clearly ν^+ and ν^- are nonnegative set functions, and also $\nu^+(\varnothing) = \nu^-(\varnothing) = 0$ is clear by the remark following Definition 2.23. In addition, the countable additivity of ν^+ and ν^- follows from Part (1), so both $(X, \mathcal{M}, \nu^+)$ and $(X, \mathcal{M}, \nu^-)$ are measure spaces by Part (2). Now, for any $A \in \mathcal{M}$,

$$\nu(A) = \int_A f\,d\mu = \int_A f^+\,d\mu - \int_A f^-\,d\mu = \nu^+(A) - \nu^-(A) = (\nu^+ - \nu^-)(A)$$

by Definition 2.22, where the existence of all integrals is assured by Theorem 2.27(5). ■

Theorem 2.31 (**Monotone Convergence Theorem**). *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $(f_n)_{n\in\mathbb{N}}$ be a monotone increasing sequence of nonnegative Borel measurable functions on $(X, \mathcal{M})$. If $f_n \uparrow f$ pointwise on X, then*

$$\lim_{n\to\infty}\int_X f_n\,d\mu = \int_X f\,d\mu.$$

Theorem 2.32 (**Additivity Theorem**). *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f_1, \ldots, f_n$ be Borel measurable functions on $(X, \mathcal{M})$ such that $\sum_{k=1}^n f_k$ is well-defined. If $\int_X f_k\,d\mu$ exists for each k and $\sum_{k=1}^n \int_X f_k\,d\mu$ is well-defined, then*

$$\int_X \sum_{k=1}^n f_k\,d\mu = \sum_{k=1}^n \int_X f_k\,d\mu.$$

Thus if $f_1, \ldots, f_n$ are μ-integrable, then so too is $\sum_{k=1}^n f_k$.

Corollary 2.33. *Let $(X, \mathcal{M}, \mu)$ be a measure space.*

1. *If $(f_n)_{n\in\mathbb{N}}$ is a sequence of nonnegative Borel measurable functions on $(X, \mathcal{M})$, then*

$$\int_X \left(\sum_{n=1}^{\infty} f_n\right) d\mu = \sum_{n=1}^{\infty} \int_X f_n \, d\mu.$$

2. *If f is Borel measurable, then f is μ-integrable if and only if $|f|$ is μ-integrable.*
3. *Suppose f and g are Borel measurable with $|f| \leq g$. If g is μ-integrable, then f is μ-integrable.*

Definition 2.34. *In a measure space $(X, \mathcal{M}, \mu)$, a property is said to hold **almost everywhere** with respect to μ (written as μ-**a.e.**) if there exists a set $A \in \mathcal{M}$ of μ-measure 0 such that the property holds on $X \setminus A$ and fails on A.*

Proposition 2.35. *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f, g : X \to \overline{\mathbb{R}}$ be Borel measurable.*

1. *If $f = 0$ μ-a.e., then $\int_X f \, d\mu = 0$.*
2. *If $f = g$ μ-a.e. and $\int_X f \, d\mu$ exists, then $\int_X g \, d\mu$ also exists with $\int_X f \, d\mu = \int_X g \, d\mu$.*
3. *If f is μ-integrable, then f is finite μ-a.e.*
4. *If $f \geq 0$ and $\int_X f \, d\mu = 0$, then $f = 0$ μ-a.e.*

Theorem 2.36 (Extended Monotone Convergence Theorem). *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of Borel measurable functions on $(X, \mathcal{M})$, and let $f, g : X \to \overline{\mathbb{R}}$ be Borel measurable.*

1. *Suppose $\int_X g \, d\mu > -\infty$. If $f_n \geq g$ for all n and $f_n \uparrow f$ pointwise on X, then $\int_X f_n \, d\mu$ exists for all n with*

$$\int_X f_n \, d\mu \uparrow \int_X f \, d\mu.$$

2. *Suppose $\int_X g \, d\mu < +\infty$. If $f_n \leq g$ for all n and $f_n \downarrow f$ pointwise on X, then $\int_X f_n \, d\mu$ exists for all n with*

$$\int_X f_n \, d\mu \downarrow \int_X f \, d\mu.$$

Theorem 2.37 (Fatou's Lemma). *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of Borel measurable functions on $(X, \mathcal{M})$.*

1. *Suppose $\int_X f \, d\mu > -\infty$. If $f_n \geq f$ for all n, then*

$$\liminf_{n\to\infty} \int_X f_n \, d\mu \geq \int_X \left(\liminf_{n\to\infty} f_n\right) d\mu.$$

2. *Suppose $\int_X f \, d\mu < +\infty$. If $f_n \leq f$ for all n, then*

$$\limsup_{n\to\infty} \int_X f_n \, d\mu \leq \int_X \left(\limsup_{n\to\infty} f_n\right) d\mu.$$

Theorem 2.38 (Dominated Convergence Theorem). *Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of Borel measurable functions on $(X, \mathcal{M})$. If $f, g : X \to \overline{\mathbb{R}}$ are Borel measurable with g μ-integrable, $|f_n| \le g$ for all n, and $f_n \to f$ pointwise μ-a.e., then f is μ-integrable and*

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Corollary 2.39. *Let $(X, \mathcal{M}, \mu)$ be a measure space, let $(f_n)_{n\in\mathbb{N}}$ be a sequence of Borel measurable functions on $(X, \mathcal{M})$, and fix $p \in (0, \infty)$. If $f, g : X \to \overline{\mathbb{R}}$ are Borel measurable with $|g|^p$ μ-integrable, $|f_n| \le g$ for all n, and $f_n \to f$ pointwise μ-a.e., then $|f|^p$ is μ-integrable and*

$$\lim_{n\to\infty} \int_X |f_n - f|^p \, d\mu = 0.$$

Proposition 2.40. *Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space, let $f, g : X \to \overline{\mathbb{R}}$ be Borel measurable, and suppose $\int_X f \, d\mu$ and $\int_X g \, d\mu$ exist. If*

$$\int_A f \, d\mu \le \int_A g \, d\mu$$

for all $A \in \mathcal{M}$, then $f \le g$ μ-a.e.

Proposition 2.41. *Suppose $(X, \mathcal{M}, \mu)$ is a measure space and $(Y, \mathcal{N})$ is a measurable space. Let $T : X \to Y$ be a $(\mathcal{M}, \mathcal{N})$-measurable mapping. Define a measure ν on $(Y, \mathcal{N})$ by $\nu(B) = \mu(T^{-1}(B))$ for all $B \in \mathcal{N}$. If $f : (Y, \mathcal{N}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ and $B \in \mathcal{N}$, then*

$$\int_{T^{-1}(B)} f \circ T \, d\mu = \int_B f \, d\nu$$

(i.e. if one of the integrals exists then so does the other, and they must be equal).

Problem 2.42 (PMT 1.6.1). Let $f : (c, d) \times (a, b) \subseteq \mathbb{R}^2 \to \mathbb{R}$ be such that $f(x, \cdot) : (a, b) \to \mathbb{R}$ is Borel measurable[6] for each $x \in (c, d)$. Assume $g : (a, b) \to \mathbb{R}$ is Borel measurable with $|f(x, y)| \le g(y)$ for all x, y, and $\int_{(a,b)} g \, d\lambda \in \mathbb{R}$ (where λ is the Lebesgue measure on $\mathbb{R}$). If $x_0 \in (c, d)$ and $\lim_{x\to x_0} f(x, y)$ exists for all $y \in (a, b)$, show that

$$\lim_{x\to x_0} \int_{(a,b)} f(x, \cdot) \, d\lambda = \int_{(a,b)} \left[\lim_{x\to x_0} f(x, \cdot) \right] d\lambda.$$

Solution. Set $I = (a, b)$, $J = (c, d)$. Since $\lim_{x\to x_0} f(x, y)$ exists for all $y \in I$, we may define $\varphi : I \to \overline{\mathbb{R}}$ by

$$\varphi(y) = \lim_{x\to x_0} f(x, y),$$

for each $y \in I$. Let $(x_n)_{n\in\mathbb{N}}$ be an arbitrary sequence in J such that $x_n \to x_0$. For each n define $f_n : I \to \mathbb{R}$ by $f_n = f(x_n, \cdot)$, so that $(f_n)_{n\in\mathbb{N}}$ is a sequence of Borel measurable functions on $(I, \mathcal{B}(I))$. Clearly

$$\varphi(y) = \lim_{n\to\infty} f_n(y)$$

[6] That is, $f(x, \cdot) : ((a, b), \mathcal{B}((a, b))) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

for each $y \in I$, and so φ is Borel measurable on $(I, \mathcal{B}(I))$ by Proposition 2.7. In addition, g is Borel measurable on $(I, \mathcal{B}(I))$ as well as λ-integrable on the measure space $(I, \mathcal{B}(I), \lambda)$, and we have $|f_n| \leq g$ for all n and $f_n \to \varphi$ pointwise λ-a.e. Therefore, by the Dominated Convergence Theorem, φ is λ-integrable on $(I, \mathcal{B}(I), \lambda)$, and

$$\lim_{n\to\infty} \int_I f(x_n, \cdot)\, d\lambda = \lim_{n\to\infty} \int_I f_n\, d\lambda = \int_I \varphi\, d\lambda = \int_I \left[\lim_{x\to x_0} f(x, \cdot)\right] d\lambda.$$

Since $(x_n)_{n\in\mathbb{N}}$ is an arbitrary sequence converging to x_0, it follows that

$$\lim_{x\to x_0} \int_I f(x, \cdot)\, d\lambda = \int_I \left[\lim_{x\to x_0} f(x, \cdot)\right] d\lambda,$$

as desired. ■

Problem 2.43 (PMT 1.6.2). Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of Borel measurable functions on the measure space $(X, \mathcal{M}, \mu)$ such that $\sum_{n=1}^{\infty} f_n(x)$ exists in $\overline{\mathbb{R}}$ for all $x \in X$. If

$$\sum_{n=1}^{\infty} \int_X |f_n|\, d\mu \in \mathbb{R},$$

show that $\sum_{n=1}^{\infty} f_n(x)$ converges in $\mathbb{R}$ μ-a.e. on X, and

$$\int_X \left(\sum_{n=1}^{\infty} f_n\right) d\mu = \sum_{n=1}^{\infty} \int_X f_n\, d\mu.$$

Solution. By Proposition 2.26, $(|f_n|)_{n\in\mathbb{N}}$ is a sequence of nonnegative Borel measurable functions. Let

$$g_n = \sum_{k=1}^{n} |f_k|$$

for each n, where each g_n is Borel measurable by Theorem 2.25. Note that $g : X \to \overline{\mathbb{R}}$ given by $g = \sum_{n=1}^{\infty} |f_n|$ is Borel measurable by Proposition 2.7 since

$$g(x) = \sum_{n=1}^{\infty} g_n(x) = \lim_{n\to\infty} \sum_{k=1}^{n} |f_k| = \lim_{n\to\infty} g_n(x).$$

By Corollary 2.30(1),

$$\int_X g\, d\mu = \int_X \left(\sum_{n=1}^{\infty} |f_n|\right) d\mu = \sum_{n=1}^{\infty} \int_X |f_n|\, d\mu \in \mathbb{R},$$

so g is μ-integrable and Proposition 2.35(3) implies that g is finite μ-a.e. Thus the series $\sum_{n=1}^{\infty} f_n(x)$ is absolutely convergent in $\mathbb{R}$ μ-a.e., and hence converges in $\mathbb{R}$ μ-a.e.

The hypothesis that $\sum_{n=1}^{\infty} f_n(x)$ exists in $\overline{\mathbb{R}}$ for all $x \in X$ implies that $\sum_{k=1}^{n} f_k(x)$ does not feature both $+\infty$ and $-\infty$ among its terms for any $n \in \mathbb{N}$ and $x \in X$, so that $\varphi_n = \sum_{k=1}^{n} f_k$ is Borel measurable for each n by Proposition 2.17(1), and moreover $\varphi : X \to \overline{\mathbb{R}}$ given by

$$\varphi(x) = \sum_{n=1}^{\infty} f_n(x) = \lim_{n\to\infty} \varphi_n(x)$$

is Borel measurable by Proposition 2.7. We now see that $(\varphi_n)_{n\in\mathbb{N}}$ is a sequence of Borel measurable functions, $\varphi, g : X \to \overline{\mathbb{R}}$ are Borel measurable with g μ-integrable, and

$$|\varphi_n| = \left|\sum_{k=1}^{n} f_n\right| \leq \sum_{k=1}^{n} |f_k| \leq \sum_{n=1}^{\infty} |f_n| = g$$

for all n, and $\varphi_n \to \varphi$ pointwise μ-a.e. By the Additivity Theorem and Dominated Convergence Theorem we conclude that

$$\begin{aligned}\sum_{n=1}^{\infty} \int_X f_n \, d\mu &= \lim_{n\to\infty} \sum_{k=1}^{n} \int_X f_k \, d\mu = \lim_{n\to\infty} \int_X \left(\sum_{k=1}^{n} f_k\right) d\mu \\ &= \lim_{n\to\infty} \int_X \varphi_n \, d\mu = \int_X \varphi \, d\mu = \int_X \left(\sum_{n=1}^{\infty} f_n\right) d\mu\end{aligned}$$

as desired. ■

For the following problem, given a function $F : S \subseteq \mathbb{R}^2 \to \mathbb{R}$, we let both F_x and $\partial F/\partial x$ denote the partial derivative of F with respect to its first argument.

Problem 2.44 (**PMT 1.6.3**). Let $I, J \subseteq \mathbb{R}$ be open intervals. Let $f : J \times I \subseteq \mathbb{R}^2 \to \mathbb{R}$ be such that $f(x, \cdot) : I \to \mathbb{R}$ is Borel measurable and $\int_I f(x, \cdot)\, d\lambda \in \mathbb{R}$ for each $x \in J$. Suppose that f_x exists on $J \times I$, and there exists Borel measurable $g : I \to \mathbb{R}$ with $\int_I g \, d\lambda \in \mathbb{R}$ such that $|f_x(x, y)| \leq g(y)$ for all $(x, y) \in J \times I$. Show that $\frac{d}{dx} \int_I f(x, \cdot)\, d\lambda$ exists in $\mathbb{R}$ for each $x \in J$, with

$$\frac{d}{dx} \int_I f(x, \cdot)\, d\lambda = \int_I \frac{\partial f}{\partial x}(x, \cdot)\, d\lambda.$$

Solution. Fix $x \in J$. By definition

$$f_x(x, y) = \lim_{h\to 0} \frac{f(x + h, y) - f(x, y)}{h}$$

for each $y \in I$. Let $(h_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R} \setminus \{0\}$ such that $h_n \to 0$ and $x + h_n \in J$ for all n. For each n define $F_n : I \to \mathbb{R}$ by

$$F_n(y) = \frac{f(x + h_n, y) - f(x, y)}{h_n}$$

for all $y \in I$. Since $f(x + h_n, \cdot), f(x, \cdot) : (I, \mathcal{B}(I)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, Theorem 2.25(2) implies that $f(x + h_n, \cdot) - f(x, \cdot)$ is Borel measurable. Also the constant function $I \to \{1/h_n\}$ is Borel measurable by Proposition 2.5, so that F_n is Borel measurable by Theorem 2.25(3). Thus $(F_n)_{n\in\mathbb{N}}$ is a sequence of Borel measurable functions such that

$$\lim_{n\to\infty} F_n(y) = f_x(x, y) \in \mathbb{R}$$

for all $y \in I$, and by Theorem 2.12 we conclude that $f_x(x, \cdot)$ is Borel measurable as well.

For each $n \in \mathbb{N}$ we have $|F_n| \leq g$. Indeed, if there were to exist some $n_0 \in \mathbb{N}$ and $y_0 \in I$ such that

$$|F_{n_0}(y_0)| = \left|\frac{f(x + h_{n_0}, y_0) - f(x, y)}{h_{n_0}}\right| > g(y_0),$$

then by the Mean Value Theorem there must be some x_0 between x and $x + h_{n_0}$ such that

$$f_x(x_0, y_0) = \frac{f(x + h_{n_0}, y_0) - f(x, y)}{h_{n_0}},$$

leading to the contradiction $|f_x(x_0, y_0)| > g(y_0)$.

Now, by the Dominated Convergence Theorem it follows that $f_x(x, \cdot)$ is λ-integrable with

$$\lim_{n\to\infty} \int_I \frac{f(x + h_n, \cdot) - f(x, \cdot)}{h_n} d\lambda = \lim_{n\to\infty} \int_I F_n \, d\lambda = \int_I f_x(x, \cdot) \, d\lambda.$$

By the arbitrariness of the sequence $(h_n)_{n\in\mathbb{N}}$ it follows that

$$\lim_{h\to 0} \int_I \frac{f(x + h, \cdot) - f(x, \cdot)}{h} d\lambda = \int_I f_x(x, \cdot) \, d\lambda \in \mathbb{R}.$$

On the other hand, by Theorems 2.27(1) and 2.32, along with the hypothesis that $\int_I f(\xi, \cdot) \, d\lambda$ exists in $\mathbb{R}$ for all $\xi \in J$, we have

$$\begin{aligned}
\lim_{h\to 0} \int_I \frac{f(x + h, \cdot) - f(x, \cdot)}{h} d\lambda &= \lim_{h\to 0} \frac{\int_I f(x + h, \cdot) - f(x, \cdot) \, d\lambda}{h} \\
&= \lim_{h\to 0} \frac{\int_I f(x + h, \cdot) \, d\lambda - \int_I f(x, \cdot) \, d\lambda}{h} \\
&= \frac{d}{dx} \int_I f(x, \cdot) \, d\lambda.
\end{aligned}$$

Since $x \in J$ is arbitrary, we are done. ■

The intervals I and J in Problem 2.44 could just as well be closed instead of open. The proof would be little altered: if $x \in J$ is an endpoint of J, then $f_x(x, y)$ would be understood to be the appropriate one-sided limit.

2.4 – The Riemann Integral

A **partition** of an interval $[a,b] \subseteq \mathbb{R}$ is a finite set of points $P = \{x_i\}_{i=0}^n \subseteq [a,b]$ such that

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

For each $1 \leq i \leq n$ we call $I_i = [x_{i-1}, x_i]$ the ith **subinterval** of the partition P, and the **length** of the i subinterval is

$$\Delta x_i = x_i - x_{i-1}.$$

The **mesh** of P is

$$\|P\| = \max_{1 \leq i \leq n} \Delta x_i.$$

The collection of all possible partitions of $[a,b]$ we denote by $\mathcal{P}([a,b])$ or $\mathcal{P}[a,b]$.[7] Next, given a function $f : [a,b] \to \mathbb{R}$ we define

$$u_i = \sup_{x \in I_i} f(x) \quad \text{and} \quad \ell_i = \inf_{x \in I_i} f(x),$$

and also

$$U(P,f) = \sum_{i=1}^n u_i \Delta x_i \quad \text{and} \quad L(P,f) = \sum_{i=1}^n \ell_i \Delta x_i.$$

We call $U(P,f)$ the **upper sum of f with respect to P** and $L(P,f)$ the **lower sum of f with respect to P**. Now we define

$$\overline{\int_a^b} f = \inf\big\{U(P,f) : P \in \mathcal{P}[a,b]\big\} \quad \text{and} \quad \underline{\int_a^b} f = \sup\big\{L(P,f) : P \in \mathcal{P}[a,b]\big\},$$

the **upper Riemann integral of f over** $[a,b]$ and **lower Riemann integral of f over** $[a,b]$, respectively. Note that if $f : [a,b] \to \mathbb{R}$ is a bounded function (i.e. $f([a,b])$ is a bounded set), then both the upper and lower Riemann integrals of f over $[a,b]$ will exist in $\mathbb{R}$. They are not always equal, however.

Definition 2.45. *If $f : [a,b] \to \mathbb{R}$ is such that*

$$\overline{\int_a^b} f = \underline{\int_a^b} f,$$

*then the **Riemann integral of f over** $[a,b]$ is defined to be*

$$\int_a^b f = \overline{\int_a^b} f.$$

*If $\int_a^b f \in \mathbb{R}$, we say f is **Riemann-integrable on** $[a,b]$. The collection of all Riemann-integrable functions on an interval $[a,b]$ is denoted by $\mathcal{R}([a,b])$ or $\mathcal{R}[a,b]$.*

[7] Recall that $\mathcal{P}([a,b])$ denotes the power set of $[a,b]$, which is the collection of all subsets of $[a,b]$.

Remark. The Lebesgue integral of a function f over an interval $[a,b]$ with respect to a measure μ will be denoted by

$$\int_{[a,b]} f\,d\mu \quad \text{or} \quad \int_{[a,b]} f(x)\,d\mu(x),$$

whereas the Riemann integral of f over $[a,b]$ will be denoted by

$$\int_a^b f \quad \text{or} \quad \int_a^b f(x)\,dx.$$

Some authors may use one or the other of the last two symbols to denote a Lebesgue integral with respect to Lebesgue measure, which is ungladsome to the utmost.

Given a partition $P \in \mathcal{P}[a,b]$, we say $P^* \in \mathcal{P}[a,b]$ is a **refinement** of P if $P^* \supseteq P$. The following proposition, and the one after it, are established in elementary analysis.

Proposition 2.46. *Let $f : [a,b] \to \mathbb{R}$. If $P^* \in \mathcal{P}[a,b]$ is a refinement of $P \in \mathcal{P}[a,b]$, then*

$$L(P,f) \le L(P^*,f) \quad \text{and} \quad U(P^*,f) \le U(P,f).$$

Given a partition $P = \{x_i\}_{i=0}^n \in \mathcal{P}[a,b]$, a **sample point** from $[x_{i-1},x_i]$ is any point $x_i^* \in [x_{i-1},x_i]$, so that $x_{i-1} \le x_i^* \le x_i$ for each $1 \le i \le n$.

Definition 2.47. *Given a function $f : [a,b] \to \mathbb{R}$, a partition $P = \{x_i\}_{i=0}^n \in \mathcal{P}[a,b]$, and sample points $x_1^*,\dots,x_n^*$, we call*

$$S(P,f) = \sum_{i=1}^n f(x_i^*)\Delta x_i$$

a ***Riemann sum of f with respect to P on*** *$[a,b]$.*

Definition 2.48. *Let $r \in \mathbb{R}$. Then we define*

$$\lim_{\|P\|\to 0} S(P,f) = r$$

to mean the following: for every $\epsilon > 0$ there exists some $\delta > 0$ such that if $P = \{x_i\}_{i=0}^n \in \mathcal{P}[a,b]$ with $\|P\| < \delta$, then

$$|S(P,f) - r| < \epsilon$$

for all choice of sample points $x_i^* \in [x_{i-1},x_i]$.

Proposition 2.49. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function. Then $f \in \mathcal{R}[a,b]$ if and only if $\lim_{\|P\|\to 0} S(P,f) \in \mathbb{R}$, in which case*

$$\int_a^b f = \lim_{\|P\|\to 0} S(P,f).$$

The **upper function** $\alpha : [a,b] \to \mathbb{R}$ corresponding to a partition $P = \{x_i\}_{i=0}^n \in \mathcal{P}[a,b]$ and function $f : [a,b] \to \mathbb{R}$ is given by

$$\alpha(x) = \begin{cases} u_1, & \text{if } x \in [x_0,x_1] \\ u_i, & \text{if } x \in (x_{i-1},x_i] \text{ for } 2 \le i \le n, \end{cases} \tag{2.10}$$

and the **lower function** $\beta : [a,b] \to \mathbb{R}$ is given by

$$\beta(x) = \begin{cases} \ell_1, & \text{if } x \in [x_0, x_1] \\ \ell_i, & \text{if } x \in (x_{i-1}, x_i] \text{ for } 2 \le i \le n. \end{cases} \tag{2.11}$$

Let $\lambda : \mathcal{B}(\mathbb{R}) \to \overline{\mathbb{R}}$ be Lebesgue measure. The σ-algebra $\mathcal{B}([a,b])$, known as the collection of Borel sets in $[a,b]$, is in fact equal to $\mathcal{B}(\mathbb{R}) \cap [a,b]$, and so we obtain the real-valued measure $\lambda : \mathcal{B}([a,b]) \to \mathbb{R}$. Define $\overline{\mathcal{B}}([a,b]) = \mathcal{B}([a,b])_\lambda$, the completion of $\mathcal{B}([a,b])$ relative to $\lambda : \mathcal{B}([a,b]) \to \mathbb{R}$.[8] Consider the measure space $([a,b], \overline{\mathcal{B}}([a,b]), \lambda)$. The functions α and β are both $(\overline{\mathcal{B}}([a,b]), \mathcal{B}(\mathbb{R}))$-measurable, which is to say Borel measurable on $([a,b], \overline{\mathcal{B}}([a,b]))$, and so both are simple functions. Moreover,

$$\int_{[a,b]} \alpha \, d\lambda = u_1 \lambda[x_0, x_1] + \sum_{i=2}^{n} u_i \lambda(x_{i-1}, x_i] = u_1 \Delta x_1 + \sum_{i=2}^{n} u_i \Delta x_i = \sum_{i=1}^{n} u_i \Delta x_i = U(P, f),$$

and similarly

$$\int_{[a,b]} \beta \, d\lambda = L(P, f).$$

Let $(P_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{P}[a,b]$ such that P_{k+1} is a refinement of P_k and $\lim_{k\to\infty} \|P_k\| = 0$. Let α_k and β_k be the upper and lower function corresponding to P_k and $f : [a,b] \to \mathbb{R}$. Since

$$\beta_1 \le \beta_k \le \beta_{k+1} \le f \le \alpha_{k+1} \le \alpha_k \le \alpha_1$$

for all $k \in \mathbb{N}$, there exist functions $\alpha, \beta : [a,b] \to \mathbb{R}$ such that $(\alpha_k)_{k\in\mathbb{N}}$ converges pointwise on $[a,b]$ to α and $(\beta_k)_{k\in\mathbb{N}}$ converges pointwise on $[a,b]$ to β. Each α_k and β_k, being a simple function, is Borel measurable on $([a,b], \overline{\mathcal{B}}([a,b]))$, and so α and β are also Borel measurable by Proposition 2.7. Suppose f is a bounded function, so there exists some $M \in \mathbb{R}$ such that $|f| \le M$. As a constant function, M is both Borel measurable and λ-integrable,

$$\int_{[a,b]} M \, d\lambda = M\lambda[a,b] = M(b-a) \in \mathbb{R},$$

and since $|\alpha_k|, |\beta_k| \le M$ for all k, the Dominated Convergence Theorem implies that α and β are λ-integrable with

$$\lim_{k\to\infty} \int_{[a,b]} \alpha_k \, d\lambda = \int_{[a,b]} \alpha \, d\lambda \quad \text{and} \quad \lim_{k\to\infty} \int_{[a,b]} \beta_k \, d\lambda = \int_{[a,b]} \beta \, d\lambda,$$

or equivalently

$$\lim_{k\to\infty} U(P_k, f) = \int_{[a,b]} \alpha \, d\lambda \quad \text{and} \quad \lim_{k\to\infty} L(P_k, f) = \int_{[a,b]} \beta \, d\lambda.$$

Lemma 2.50. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function. Let $(P_k)_{k\in\mathbb{N}} \subseteq \mathcal{P}[a,b]$ with $P_k \subseteq P_{k+1}$ and $\|P_k\| \to 0$, and let $\alpha, \beta : [a,b] \to \mathbb{R}$ be as defined by* (2.10) *and* (2.11)*. If $x \notin \bigcup_{k=1}^{\infty} P_k$, then f is continuous at x if and only if $\alpha(x) = f(x) = \beta(x)$.*

The results above, in addition to the following proposition, will be needed to prove Theorem 2.52 below.

[8]See Example 1.52. We call $\overline{\mathcal{B}}([a,b])$ the collection of Lebesgue measurable sets in $[a,b]$.

Proposition 2.51. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function. Then $f \in \mathcal{R}[a,b]$ if and only if there exists some $r \in \mathbb{R}$ such that*

$$\lim_{k\to\infty} U(P_k, f) = \lim_{k\to\infty} L(P_k, f) = r \tag{2.12}$$

for all $(P_k)_{k\in\mathbb{N}} \subseteq \mathcal{P}[a,b]$ with $P_k \subseteq P_{k+1}$ and $\|P_k\| \to 0$.

Proof. Suppose that $f \in \mathcal{R}[a,b]$, so $\int_a^b f = r$ for some $r \in \mathbb{R}$. Let $(P_k)_{k\in\mathbb{N}} \subseteq \mathcal{P}[a,b]$ with $P_k \subseteq P_{k+1}$ and $\|P_k\| \to 0$. Let $\epsilon > 0$. By Proposition 2.49 there exists some $\delta > 0$ such that, for all $P \in \mathcal{P}[a,b]$ with $\|P\| < \delta$, we have

$$|S(P,f) - r| = \left| S(P,f) - \int_a^b f \right| < \frac{\epsilon}{2}$$

for all possible choices of sample points. Let $k_0 \in \mathbb{N}$ be such that $\|P_k\| < \delta$ for all $k \geq k_0$. We have $P_{k_0} = \{x_i\}_{i=0}^n$ for some $n \in \mathbb{N}$. For each $1 \leq i \leq n$ choose $x_i^* \in [x_{i-1}, x_i]$ such that

$$0 \leq u_i - f(x_i^*) < \frac{\epsilon}{2(b-a)}.$$

Now,

$$\begin{aligned} |U(P_{k_0}, f) - S(P_{k_0}, f)| &= \left| \sum_{i=1}^n \big(u_i - f(x_i^*)\big)\Delta x_i \right| \leq \sum_{i=1}^n \big|u_i - f(x_i^*)\big|\Delta x_i \\ &< \sum_{i=1}^n \frac{\epsilon}{2(b-a)}\Delta x_i = \frac{\epsilon}{2(b-a)} \sum_{i=1}^n \Delta x_i = \frac{\epsilon}{2}, \end{aligned}$$

and so

$$|U(P_{k_0}, f) - r| \leq |U(P_{k_0}, f) - S(P_{k_0}, f)| + |S(P_{k_0}, f) - r| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

By Proposition 2.46, Definition 2.45, and the definition of the upper Riemann integral we have

$$r \leq U(P_k, f) \leq U(P_{k_0}, f)$$

for all $k \geq k_0$, implying that $|U(P_k, f) - r| < \epsilon$ for all $k \geq k_0$. Hence $\lim_{k\to\infty} U(P_k, f) = r$. The proof that $\lim_{k\to\infty} L(P_k, f) = r$ is similar.

For the converse, suppose there exists $r \in \mathbb{R}$ such that (2.12) holds for all $(P_k)_{k\in\mathbb{N}} \subseteq \mathcal{P}[a,b]$ with $P_k \subseteq P_{k+1}$ and $\|P_k\| \to 0$. Suppose

$$\rho := \overline{\int_a^b} f < r.$$

Now, for each $k \in \mathbb{N}$ there exists $Q_k \in \mathcal{P}[a,b]$ such that $\rho \leq U(Q_k, f) < \rho + 1/k$. We may assume $\|Q_k\| < 1/k$ since it is always possible, if necessary, to include the additional points

$$x_i = a + \frac{i}{k}(b-a), \quad 1 \leq i \leq k-1$$

in the partition. For each $k \in \mathbb{N}$ define

$$P_k = \bigcup_{j=1}^k Q_j.$$

Then $(P_k)_{k\in\mathbb{N}}$ is a sequence in $\mathcal{P}[a,b]$ such that $\|P_k\| \to 0$, and since $\rho \le U(P_k, f) < \rho + 1/k$ by Proposition 2.46 it is clear that $\lim_{k\to\infty} U(P_k, f) = \rho < r$, which is a contradiction. If $\rho > r$, then upon constructing a sequence $(P_k)_{k\in\mathbb{N}}$ with $P_k \subseteq P_{k+1}$ and $\|P_k\| \to 0$, we find immediately that $\lim_{k\to\infty} U(P_k, f) \ge \rho > r$, again a contradiction. Hence $\rho = r$, and a similar argument shows that

$$\underline{\int_a^b} f = r.$$

Thus $\int_a^b f = r \in \mathbb{R}$, and therefore $f \in \mathcal{R}[a,b]$. ∎

Theorem 2.52. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function.*

1. *$f \in \mathcal{R}[a,b]$ if and only if f is continuous λ-a.e. on $[a,b]$.*
2. *If $f \in \mathcal{R}[a,b]$, then f is λ-integrable on $[a,b]$ and*

$$\int_a^b f = \int_{[a,b]} f \, d\lambda.$$

We now give a brief summary of a few kinds of **improper Riemann integrals**, a topic covered in greater detail in elementary analysis. If $f \in \mathcal{R}[a,t]$ for all $t \ge a$, then we define

$$\int_a^\infty f = \lim_{t\to\infty} \int_a^t f$$

and say $\int_a^\infty f$ **converges** if $\int_a^\infty f = \alpha$ for some $\alpha \in \mathbb{R}$. If $f \in \mathcal{R}[t,b]$ for all $t \le b$, then we define

$$\int_{-\infty}^b f = \lim_{t\to-\infty} \int_t^b f$$

and say $\int_{-\infty}^b f$ **converges** if $\int_{-\infty}^b f = \beta$ for some $\beta \in \mathbb{R}$. Finally, if $f \in \mathcal{R}[s,t]$ for all $-\infty < s < t < \infty$, and $\int_{-\infty}^c f$ and $\int_c^\infty f$ both converge for some $c \in \mathbb{R}$, then we define

$$\int_{-\infty}^\infty f = \int_{-\infty}^c f + \int_c^\infty f.$$

and say $\int_{-\infty}^\infty f$ **converges**. The value of $\int_{-\infty}^\infty f$, of course, does not depend on the choice for c. Any improper Riemann integral that does not converge to a real number is said to **diverge**.

Theorem 2.53. *Let $f : \mathbb{R} \to \mathbb{R}$ be such that $\int_{-\infty}^\infty f$ converges. Then the following hold.*

1. *The function f is continuous λ-a.e. on $\mathbb{R}$.*
2. *If f is nonnegative, then f is λ-integrable on $(\mathbb{R}, \overline{\mathcal{B}}(\mathbb{R}), \lambda)$ with*

$$\int_{\mathbb{R}} f \, d\lambda = \int_{-\infty}^\infty f.$$

Proposition 2.54. *If $f : \mathbb{R} \to \mathbb{R}$ is such that $\int_{-\infty}^\infty f^+$ and $\int_{-\infty}^\infty f^-$ both converge, then f is λ-integrable on $(\mathbb{R}, \overline{\mathcal{B}}(\mathbb{R}), \lambda)$ with*

$$\int_{\mathbb{R}} f \, d\lambda = \int_{-\infty}^\infty f.$$

Proof. Suppose $f : \mathbb{R} \to \mathbb{R}$ is such that $\int_{-\infty}^{\infty} f^+$ and $\int_{-\infty}^{\infty} f^-$ converge. Since f^+ and f^- are nonnegative, by Theorem 2.53 it follows that both functions are λ-integrable (and hence Borel measurable) on the measure space $(\mathbb{R}, \overline{\mathcal{B}}(\mathbb{R}), \lambda)$ with

$$\int_{\mathbb{R}} f^+ \, d\lambda = \int_{-\infty}^{\infty} f^+ \in \mathbb{R} \quad \text{and} \quad \int_{\mathbb{R}} f^- \, d\lambda = \int_{-\infty}^{\infty} f^- \in \mathbb{R}.$$

Now, $f = f^+ - f^-$, so f is Borel measurable by Theorem 2.25, and then by Definition 2.22 and a limit law we have

$$\int_{\mathbb{R}} f \, d\lambda = \int_{\mathbb{R}} f^+ \, d\lambda - \int_{\mathbb{R}} f^- \, d\lambda = \int_{-\infty}^{\infty} f^+ - \int_{-\infty}^{\infty} f^- = \int_{-\infty}^{\infty} (f^+ - f^-) = \int_{-\infty}^{\infty} f.$$

This makes clear that $\int_{-\infty}^{\infty} f$ is real-valued, and so f is λ-integrable on $(\mathbb{R}, \overline{\mathcal{B}}(\mathbb{R}), \lambda)$. ■

2.5 – Integration of Complex-Valued Functions

We now consider complex-valued functions on a measurable space $(X, \mathcal{M})$. If $f : X \to \mathbb{C}$, it is convenient to write

$$f(x) = u(x) + iv(x),$$

where $u : X \to \mathbb{R}$ and $v : X \to \mathbb{R}$ are the **real** and **imaginary** parts of f, respectively. We define $\operatorname{Re} f = u$ and $\operatorname{Im} f = v$, so that

$$(\operatorname{Re} f)(x) = u(x) = \operatorname{Re}[f(x)] \quad \text{and} \quad (\operatorname{Im} f)(x) = v(x) = \operatorname{Im}[f(x)] \tag{2.13}$$

The symbols $\operatorname{Re} f(x)$ and $\operatorname{Im} f(x)$ may be used instead of $(\operatorname{Re} f)(x)$ and $(\operatorname{Im} f)(x)$.

Definition 2.55. *Let $(X, \mathcal{M})$ be a measurable space. We say $f : X \to \mathbb{C}$ is **Borel measurable** if $\operatorname{Re} f : X \to \mathbb{R}$ and $\operatorname{Im} f : X \to \mathbb{R}$ are both Borel measurable. If μ is a measure on $\mathcal{M}$, and both $\operatorname{Re} f$ and $\operatorname{Im} f$ are μ-integrable, then we define*

$$\int_X f\,d\mu = \int_X \operatorname{Re} f\,d\mu + i\int_X \operatorname{Im} f\,d\mu$$

*and say that f is μ-**integrable**.*

If f is a complex-valued function we do not entertain the integral $\int_X f\,d\mu$ when either $\int_X \operatorname{Re} f\,d\mu$ or $\int_X \operatorname{Im} f\,d\mu$ is not real-valued.

The Riemann integral constructed in the previous section also has a natural extension to complex-valued functions. The **Riemann integral** of $f : [a, b] \to \mathbb{C}$ is defined to be

$$\int_a^b f = \int_a^b \operatorname{Re} f + i\int_a^b \operatorname{Im} f, \tag{2.14}$$

provided the integrals on the right-hand side exist in $\mathbb{R}$, in which case we say φ is **Riemann integrable** on $[a, b]$. By letting $a = -\infty$ or $b = \infty$, we may extend the definition (2.14) to include some improper Riemann integrals. In particular, given $f : \mathbb{R} \to \mathbb{C}$, we define

$$\int_{-\infty}^{\infty} f = \int_{-\infty}^{\infty} \operatorname{Re} f + i\int_{-\infty}^{\infty} \operatorname{Im} f \tag{2.15}$$

provided $\int_{-\infty}^{\infty} \operatorname{Re} f \in \mathbb{R}$ and $\int_{-\infty}^{\infty} \operatorname{Im} f \in \mathbb{R}$.

Proposition 2.56. *If $f : \mathbb{R} \to \mathbb{C}$ is such that $\int_{-\infty}^{\infty} |f|$ converges, then f is λ-integrable on $(\mathbb{R}, \overline{\mathcal{B}}(\mathbb{R}), \lambda)$ with*

$$\int_{\mathbb{R}} f\,d\lambda = \int_{-\infty}^{\infty} f.$$

Proof. Suppose $f : \mathbb{R} \to \mathbb{C}$ is such that $\int_{-\infty}^{\infty} |f| \in \mathbb{R}$. Let $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$. Since

$$|f| = \sqrt{u^2 + v^2}, \quad |u| = u^+ + u^-, \quad \text{and} \quad |v| = v^+ + v^-,$$

we have

$$0 \le u^-, u^+ \le |u| \le |f| \quad \text{and} \quad 0 \le v^-, v^+ \le |v| \le |f|,$$

and so by the Comparison Test for Integrals from calculus it follows that

$$\int_{-\infty}^{\infty} u^+, \quad \int_{-\infty}^{\infty} u^-, \quad \int_{-\infty}^{\infty} v^+, \quad \text{and} \quad \int_{-\infty}^{\infty} v^-$$

all converge, and therefore

$$\int_{-\infty}^{\infty} u = \int_{\mathbb{R}} u \, d\lambda \in \mathbb{R} \quad \text{and} \quad \int_{-\infty}^{\infty} v = \int_{\mathbb{R}} v \, d\lambda \in \mathbb{R}$$

by Proposition 2.54. Now

$$\int_{\mathbb{R}} f \, d\lambda = \int_{\mathbb{R}} u \, d\lambda + i \int_{\mathbb{R}} v \, d\lambda = \int_{-\infty}^{\infty} u + i \int_{-\infty}^{\infty} v = \int_{-\infty}^{\infty} f$$

by Definition 2.55 and (2.15). ■

3

Product Spaces

3.1 – Product σ-Algebras

Definition 3.1. *Let $(X_1, \mathcal{M}_1), \ldots, (X_n, \mathcal{M}_n)$ be measurable spaces, and set $X = X_1 \times \cdots \times X_n$. The set of* ***measurable rectangles in*** *X is the collection $\mathcal{R}(X) \subseteq \mathcal{P}(X)$ given by*

$$\mathcal{R}(X) = \big\{A_1 \times \cdots \times A_n : \forall 1 \leq k \leq n(A_k \in \mathcal{M}_k)\big\},$$

and the ***product σ-algebra on*** *X is*

$$\mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n = \sigma_X(\mathcal{R}(X)) = \bigcap\{\mathcal{S} \supseteq \mathcal{R}(X) : \mathcal{S} \text{ is a } \sigma\text{-algebra on } X\}.$$

We may also denote $\mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n$ by $\bigotimes_{k=1}^n \mathcal{M}_k$. If $\mathcal{M}_k = \mathcal{M}$ for all $1 \leq k \leq n$, then we define $\mathcal{M}^{\otimes n} = \mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n$.

The following lemma gives some properties of projection maps $\Pi_k : X_1 \times \cdots \times X_n \to X_k$, first defined by (2.1), which are routine to verify.

Lemma 3.2. *Let $X_1, \ldots, X_n$ be nonempty sets, and let $\Pi_k : X_1 \times \cdots \times X_n \to X_k$ be the kth projection map: $\Pi_k(\mathbf{x}) = [\mathbf{x}]_k$.*

1. *For any $1 \leq k \leq n$ and $S \subseteq X_k$,*

$$\Pi_k^{-1}(S) = X_1 \times \cdots \times X_{k-1} \times S \times X_{k+1} \times \cdots \times X_n$$

and

$$(X_1 \times \cdots \times X_n) \setminus \Pi_k^{-1}(S) = \Pi_k^{-1}(X_k \setminus S).$$

2. *If $S_k \subseteq X_k$ for each $1 \leq k \leq n$, then*

$$\bigcap_{k=1}^n \Pi_k^{-1}(S_k) = \prod_{k=1}^n S_k.$$

Proposition 3.3. *For each $1 \leq k \leq n$ let $(X_k, \mathcal{M}_k)$ be a measurable space, and let $\{X_k\} \subseteq \mathcal{E}_k \subseteq \mathcal{P}(X_k)$ be such that $\sigma_{X_k}(\mathcal{E}_k) = \mathcal{M}_k$. Set $X = X_1 \times \cdots \times X_n$. If*

$$\mathcal{E} = \big\{E_1 \times \cdots \times E_n : \forall 1 \leq k \leq n(E_k \in \mathcal{E}_k)\big\},$$

then $\sigma_X(\mathcal{E}) = \bigotimes_{k=1}^n \mathcal{M}_k$.

Proof. Since $\mathcal{E}_k \subseteq \mathcal{M}_k$ for each $1 \leq k \leq n$, it is clear that $\mathcal{E} \subseteq \mathcal{R}(X)$, and hence $\sigma_X(\mathcal{E}) \subseteq \bigotimes_{k=1}^n \mathcal{M}_k$.

For each $1 \leq k \leq n$ define

$$\mathcal{F}_k = \left\{F \subseteq X_k : \Pi_k^{-1}(F) \in \sigma_X(\mathcal{E})\right\}.$$

Since $X_j \in \mathcal{E}_j$ for each $1 \leq j \leq n$, for any $E \in \mathcal{E}_k$ we have

$$\Pi_k^{-1}(E) = X_1 \times \cdots \times X_{k-1} \times E \times X_{k+1} \times \cdots \times X_n \in \mathcal{E} \subseteq \sigma_X(\mathcal{E}),$$

and so $\mathcal{E}_k \subseteq \mathcal{F}_k$. In particular $X_k \in \mathcal{F}_k$. If $(F_j)_{j\in\mathbb{N}}$ is a sequence in $\mathcal{F}_k$, so that $\Pi_k^{-1}(F_j) \in \sigma_X(\mathcal{E})$ for each j, then

$$\Pi_k^{-1}\left(\bigcup_{j=1}^n F_j\right) = \bigcup_{j=1}^n \Pi_k^{-1}(F_j) \in \sigma_X(\mathcal{E})$$

shows that $\bigcup_{j=1}^n F_j \in \mathcal{F}_k$ as well, and so $\mathcal{F}_k$ is closed under countable unions. By Lemma 3.2(1),

$$F \in \mathcal{F}_k \quad\Leftrightarrow\quad \Pi_k^{-1}(F) \in \sigma_X(\mathcal{E}) \quad\Leftrightarrow\quad \Pi_k^{-1}(X_k \setminus F) = X \setminus \Pi_k^{-1}(F) \in \sigma_X(\mathcal{E}) \quad\Leftrightarrow\quad X_k \setminus F \in \mathcal{F}_k,$$

and so $\mathcal{F}_k$ is closed under complementation. Thus $\mathcal{F}_k$ is a σ-algebra containing $\mathcal{E}_k$, and therefore $\mathcal{M}_k = \sigma_{X_k}(\mathcal{E}_k) \subseteq \mathcal{F}_k$.

Finally, let $A_k \in \mathcal{M}_k$ for each k. Applying Lemma 3.2(2),

$$\begin{aligned} \forall k\big(A_k \in \mathcal{M}_k\big) \quad&\Rightarrow\quad \forall k\big(A_k \in \mathcal{F}_k\big) \quad\Rightarrow\quad \forall k\big(\Pi_k^{-1}(A_k) \in \sigma_X(\mathcal{E})\big) \\ &\Rightarrow\quad \prod_{k=1}^n A_k = \bigcap_{k=1}^n \Pi_k^{-1}(A_k) \in \sigma_X(\mathcal{E}), \end{aligned}$$

which shows that $\mathcal{R}(X) \subseteq \sigma_X(\mathcal{E})$, and thus $\bigotimes_{k=1}^n \mathcal{M}_k \subseteq \sigma_X(\mathcal{E})$. ■

Corollary 3.4. *For all $n \in \mathbb{N}$, $\mathcal{B}(\mathbb{R}^n) = \bigotimes_{k=1}^n \mathcal{B}(\mathbb{R})$.*

Proof. For each $1 \leq k \leq n$ let $(X_k, \mathcal{M}_k) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and let $\mathcal{E}_k = \mathcal{O}$, where $\{\mathbb{R}\} \subseteq \mathcal{O} \subseteq \mathcal{P}(\mathbb{R})$ is the collection of all open intervals in $\mathbb{R}$. In Example 1.12 we found that $\sigma_{\mathbb{R}}(\mathcal{O}) = \mathcal{B}(\mathbb{R})$. Setting

$$\mathcal{E} = \left\{I_1 \times \cdots \times I_n : \forall 1 \leq k \leq n(I_k \in \mathcal{O})\right\},$$

by Proposition 3.3 it follows that $\sigma_{\mathbb{R}^n}(\mathcal{E}) = \bigotimes_{k=1}^n \mathcal{B}(\mathbb{R})$. However it is also known that $\sigma_{\mathbb{R}^n}(\mathcal{E}) = \mathcal{B}(\mathbb{R}^n)$ since $\mathcal{E}$ is a basis for the standard topology on $\mathbb{R}^n$, and so the proof is done. ■

Proposition 3.5. *Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces. If $A \in \mathcal{M}$ and $B \in \mathcal{N}$, then*

$$(\mathcal{M} \cap A) \otimes (\mathcal{N} \cap B) = (\mathcal{M} \otimes \mathcal{N}) \cap (A \times B).$$

Proof. Noting that $(A, \mathcal{M} \cap A)$ and $(B, \mathcal{N} \cap B)$ are measurable spaces, we set

$$\mathcal{R}(A \times B) = \left\{E \times F : E \in \mathcal{M} \cap A \text{ and } F \in \mathcal{N} \cap B\right\}$$

so that by Definition 3.1 we have

$$(\mathcal{M} \cap A) \otimes (\mathcal{N} \cap B) = \sigma_{A\times B}(\mathcal{R}(A \times B)).$$

Now, $\Omega \in \mathcal{R}(A \times B)$ implies $\Omega = (M \cap A) \times (N \cap B)$ for some $M \in \mathcal{M}$ and $N \in \mathcal{N}$, and since

$$(M \cap A) \times (N \cap B) = (M \times N) \cap (A \times B) \in \mathcal{R}(X \times Y) \cap (A \times B),$$

we see that $\mathcal{R}(A \times B) \subseteq \mathcal{R}(X \times Y) \cap (A \times B)$. Reversing the argument reverses the containment, and so

$$\mathcal{R}(A \times B) = \mathcal{R}(X \times Y) \cap (A \times B).$$

Applying Proposition 1.15, we thus obtain

$$\begin{aligned}(\mathcal{M} \cap A) \otimes (\mathcal{N} \cap B) &= \sigma_{A\times B}(\mathcal{R}(A \times B)) = \sigma_{A\times B}\big(\mathcal{R}(X \times Y) \cap (A \times B)\big)\\ &= \sigma_{X\times Y}\big(\mathcal{R}(X \times Y)\big) \cap (A \times B) = (\mathcal{M} \otimes \mathcal{N}) \cap (A \times B)\end{aligned}$$

as desired. ■

For nonempty sets X and Y let $S \subseteq X \times Y$. For $x \in X$ and $y \in Y$, define the ***x*-section** $S_x \subseteq Y$ and ***y*-section** $S^y \subseteq X$ of S to be the sets

$$S_x = \{y \in Y : (x, y) \in S\} \quad \text{and} \quad S^y = \{x \in X : (x, y) \in S\}.$$

Proposition 3.6. *Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces. If $S \in \mathcal{M} \otimes \mathcal{N}$, then $S_x \in \mathcal{N}$ for all $x \in X$ and $S^y \in \mathcal{M}$ for all $y \in Y$.*

Proof. Define the collection

$$\mathcal{S} = \big\{S \in \mathcal{M} \otimes \mathcal{N} : \forall x \in X\,(S_x \in \mathcal{N}) \text{ and } \forall y \in Y\,(S^y \in \mathcal{M})\big\}.$$

If $A \in \mathcal{M}$ and $B \in \mathcal{N}$, then

$$(A \times B)_x = \begin{cases} B, & x \in A \\ \varnothing & x \notin A \end{cases}$$

shows that $(A \times B)_x \in \mathcal{N}$ for all $x \in X$. Similarly $(A \times B)^y \in \{A, \varnothing\} \subseteq \mathcal{M}$ for all $y \in Y$, and so $A \times B \in \mathcal{S}$. That is, $\mathcal{R}(X \times Y) \subseteq \mathcal{S}$, which also makes clear that $X \times Y \in \mathcal{S}$.

Suppose $S \in \mathcal{S}$. Fix $x \in X$. We have

$$(S^c)_x = \{y \in Y : (x, y) \in S^c\} = \{y \in Y : (x, y) \notin S\} = (S_x)^c \in \mathcal{N}$$

since $S_x \in \mathcal{N}$. Similarly $(S^c)^y = (S^y)^c \in \mathcal{M}$ for all $y \in Y$, and we conclude that $S^c \in \mathcal{S}$.

Next, let $(S_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{S}$. Thus $S_n \in \mathcal{M} \otimes \mathcal{N}$ for each n, with $(S_n)_x \in \mathcal{N}$ and $(S_n)^y \in \mathcal{M}$ for each $x \in X$ and $y \in Y$. Fix $x \in X$. Now,

$$\begin{aligned} y \in \left(\bigcup_{n=1}^{\infty} S_n\right)_x \;&\Leftrightarrow\; (x, y) \in \bigcup_{n=1}^{\infty} S_n \;\Leftrightarrow\; \exists k \in \mathbb{N}\big[(x, y) \in S_k\big]\\ &\Leftrightarrow\; \exists k \in \mathbb{N}\big[y \in (S_k)_x\big] \;\Leftrightarrow\; y \in \bigcup_{n=1}^{\infty} (S_n)_x, \end{aligned}$$

and so $\big(\bigcup_{n=1}^{\infty} S_n\big)_x = \bigcup_{n=1}^{\infty} (S_n)_x \in \mathcal{N}$. Similarly $\big(\bigcup_{n=1}^{\infty} S_n\big)^y = \bigcup_{n=1}^{\infty} (S_n)^y \in \mathcal{M}$ for any $y \in Y$. Since $\bigcup_{n=1}^{\infty} S_n \in \mathcal{M} \otimes \mathcal{N}$, it follows that $\bigcup_{n=1}^{\infty} S_n \in \mathcal{S}$.

Therefore $\mathcal{S}$ is a σ-algebra containing $\mathcal{R}(X \times Y)$, which implies that $\mathcal{S}$ contains $\mathcal{M} \otimes \mathcal{N} = \sigma_{X\times Y}(\mathcal{R}(X \times Y))$. The conclusion of the proposition follows. ■

Proposition 3.7. *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ-finite measure spaces. For each $S \in \mathcal{M} \otimes \mathcal{N}$, the functions $f_S : X \to \overline{\mathbb{R}}$ and $g_S : Y \to \overline{\mathbb{R}}$ given by $f_S(x) = \nu(S_x)$ and $g_S(y) = \mu(S^y)$ are Borel measurable.*

Proof. First we note that, for any $S \in \mathcal{M} \otimes \mathcal{N}$, the function f_S is well-defined on X since $S_x \in \mathcal{N}$ for all $x \in X$ by Proposition 3.6. Assume $(Y, \mathcal{N}, \nu)$ is a finite measure space, let

$$\mathcal{S} = \big\{S \in \mathcal{M} \otimes \mathcal{N} : f_S \text{ is Borel measurable}\big\},$$

and let $(S_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{S}$ such that $S_n \uparrow S$. Then each $f_n := f_{S_n}$ is a (real-valued) Borel measurable function. Define $f : X \to \mathbb{R}$ by

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \nu((S_n)_x), \tag{3.1}$$

which exists in $\mathbb{R}$ for each $x \in X$ since $(S_n)_x \subseteq (S_{n+1})_x \subseteq S_x$ for all n implies

$$f_n(x) \le f_{n+1}(x) \le \nu(S_x) < \infty,$$

and so f is Borel measurable by Proposition 2.7. On the other hand, $((S_n)_x)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{N}$ such that $(S_n)_x \uparrow S_x$, so by Theorem 1.21(1),

$$f(x) = \lim_{n\to\infty} \nu((S_n)_x) = \nu(S_x) = f_S(x),$$

and hence f_S is Borel measurable. This shows that $S \in \mathcal{S}$ whenever S is the limit of an increasing sequence in $\mathcal{S}$.

Next suppose $(S_n)_{n\in\mathbb{N}} \subseteq \mathcal{S}$ is such that $S_n \downarrow S$. Again defining f by (3.1), only now noting that $f_n(x) \ge f_{n+1}(x)$ for all x and n, Proposition 2.7 once more implies that f is Borel measurable. But $f_S = f$ by Theorem 1.21(2), our assumption that ν is finite ensuring that $\nu(S_1) \in \mathbb{R}$, and so f_S is Borel measurable. This shows that $S \in \mathcal{S}$ whenever S is the limit of a decreasing sequence in $\mathcal{S}$, and therefore $\mathcal{S}$ is a monotone class.

The collection $\mathcal{R} := \mathcal{R}(X \times Y)$ is clearly an elementary family on $X \times Y$, and so the collection $\mathcal{A}$ of all finite disjoint unions of elements of $\mathcal{R}$ is an algebra on $X \times Y$ by Proposition 1.9. Let $A \in \mathcal{A}$, so there exist $E_1, \dots, E_n \in \mathcal{M}$ and $F_1, \dots, F_n \in \mathcal{N}$ such that $A = \bigsqcup_{k=1}^n (E_k \times F_k)$. Now,

$$A_x = \big\{y \in Y : (x,y) \in A\big\} = \bigsqcup_{k=1}^n \big\{y \in Y : (x,y) \in E_k \times F_k\big\} = \bigsqcup_{k=1}^n (E_k \times F_k)_x,$$

where $(E_k \times F_k)_x = F_k$ if $x \in E_k$, and $(E_k \times F_k)_x = \varnothing$ if $x \notin E_k$. Thus $f_A : X \to \mathbb{R}$ is given by

$$f_A(x) = \nu(A_x) = \sum_{k=1}^n \nu\big((E_k \times F_k)_x\big) = \sum_{k=1}^n \nu(F_k)\chi_{E_k}(x).$$

Since $E_k \in \mathcal{M}$, Proposition 2.15 implies $\chi_{E_k} : X \to \mathbb{R}$ is Borel measurable, and hence f_A is Borel measurable by Proposition 2.5 and Theorem 2.25(1,3). It follows that $A \in \mathcal{S}$, so $\mathcal{A} \subseteq \mathcal{S}$ and by the Monotone Class Theorem we have $\sigma(\mathcal{A}) \subseteq \mathcal{S}$. Finally, $\mathcal{R} \subseteq \mathcal{A}$ implies $\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{R}) \subseteq \mathcal{S}$, and therefore $f_S : X \to \mathbb{R}$ is Borel measurable for all $S \in \mathcal{M} \otimes \mathcal{N}$ if ν is a finite measure.

Now suppose that $(Y, \mathcal{N}, \nu)$ is σ-finite. Let $(Y_n)_{n\in\mathbb{N}} \subseteq \mathcal{N}$ be such that $Y = \bigcup_{n=1}^\infty Y_n$ and $\nu(Y_n) < \infty$ for each n. Let $S \in \mathcal{M} \otimes \mathcal{N}$, and for each $k \ge 1$ define $S_k = S \cap (X \times B_k)$, where $B_k := \bigcup_{n=1}^k Y_n \in \mathcal{N}$. Proposition 3.5 implies $S_k \in \mathcal{M} \otimes (\mathcal{N} \cap B_k)$, and since $(B_k, \mathcal{N} \cap B_k, \nu)$ is a finite measure space, it follows by our earlier argument that $f_{S_k} : X \to \overline{\mathbb{R}}$ is Borel measurable. For each $x \in X$ define

$$f(x) = \lim_{k\to\infty} f_{s_k}(x) = \lim_{k\to\infty} \nu\big((S_k)_x\big).$$

The sequence $(\nu\big((S_k)_x)\big)_{k\in\mathbb{N}}$ is monotone increasing in $[0, \infty)$ for each x, so that the limit exists in $\overline{\mathbb{R}}$ and thus $f : X \to \overline{\mathbb{R}}$ is Borel measurable by Proposition 2.7. Finally, $(S_k)_x \to S_x$ as $k \to \infty$, so by Theorem 1.21(1) we have $f(x) = \nu(S_x) = f_S(x)$, and therefore f_S is Borel measurable. The proof that g_S is Borel measurable for each $S \in \mathcal{M} \otimes \mathcal{N}$ is similar. ∎

3.2 – Product Measure Theorem

A bit of new notation is in order to help ensure that the statements of upcoming results are as clear as can be. If $(X, \mathcal{M}, \mu)$ is a measure space and $f : X \to \overline{\mathbb{R}}$ is Borel measurable, then we define

$$\int_X f(x)\,\mu(dx) = \int_X f\,d\mu.$$

Theorem 3.8 (**Product Measure Theorem**). *Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space, $(Y, \mathcal{N})$ a measurable space, and*

$$\mathcal{F} = \left\{ \nu(x, \cdot) : \mathcal{N} \to \overline{\mathbb{R}} \;\middle|\; x \in X \right\} \tag{3.2}$$

a family of set functions with the following properties:

P1. *$\nu(x, \cdot)$ is a measure on $\mathcal{N}$ for each $x \in X$.*
P2. *$\nu(\cdot, B) : X \to \overline{\mathbb{R}}$ is Borel measurable for each $B \in \mathcal{N}$.*
P3. *$\mathcal{F}$ is* ***uniformly σ-finite****: There exist sequences $(B_n)_{n\in\mathbb{N}} \subseteq \mathcal{N}$ and $(c_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}_+$ such that $Y = \bigcup_{n=1}^{\infty} B_n$ and $\nu(\cdot, B_n) \le c_n$ for each n.*

Then the function $x \mapsto \nu(x, S_x)$ is Borel measurable for each $S \in \mathcal{M} \otimes \mathcal{N}$, and $\pi : \mathcal{M} \otimes \mathcal{N} \to \overline{\mathbb{R}}$ given by

$$\pi(S) = \int_X \nu(x, S_x)\,\mu(dx)$$

is the unique measure for which

$$\pi(A \times B) = \int_A \nu(x, B)\,\mu(dx)$$

for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$.

The set function $\pi : \mathcal{M} \otimes \mathcal{N} \to \overline{\mathbb{R}}$ defined in Theorem 3.8 is a **product measure** on $\mathcal{M} \otimes \mathcal{N}$, and $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \pi)$ is a **product space**.

Corollary 3.9 (**Classical Product Measure Theorem**). *If $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces, then the set function $\mu \times \nu : \mathcal{M} \otimes \mathcal{N} \to \overline{\mathbb{R}}$ given by*

$$(\mu \times \nu)(S) = \int_X \nu(S_x)\,\mu(dx) = \int_Y \mu(S^y)\,\nu(dy)$$

is the unique measure on $\mathcal{M} \otimes \mathcal{N}$ for which

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$$

for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$.

Proof. For each $x \in X$ let $\nu(x, \cdot) = \nu$, so the family of set functions (3.2) is simply $\mathcal{F} = \{\nu\}$, which clearly satisfies property P1 in Theorem 3.8. For any $B \in \mathcal{N}$ we have $\nu(x, B) = \nu(B)$ for all $x \in X$, a constant function and hence Borel measurable, showing property P2 to be satisfied. Finally, the σ-finiteness of $(Y, \mathcal{N}, \nu)$ ensures that $\mathcal{F} = \{\nu\}$ possesses property P3. By the Product Measure Theorem it follows that the set function $\pi : \mathcal{M} \otimes \mathcal{N} \to \overline{\mathbb{R}}$ given by

$$\pi(S) = \int_X \nu(S_x)\,\mu(dx)$$

is the unique measure for which

$$\pi(A \times B) = \int_A \nu(x, B)\, \mu(dx) = \int_A \nu(B)\, d\mu = \mu(A)\nu(B)$$

for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$.

Now, by a symmetrical argument that reverses the roles of $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$, the Product Measure Theorem implies the set function $\hat{\pi} : \mathcal{M} \otimes \mathcal{N} \to \overline{\mathbb{R}}$ given by

$$\hat{\pi}(S) = \int_Y \mu(S^y)\, \nu(dy)$$

is the unique measure for which

$$\hat{\pi}(A \times B) = \int_B \mu(A, y)\, \nu(dy) = \int_B \mu(A)\, d\nu = \mu(A)\nu(B)$$

for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$. By uniqueness it follows that $\pi = \hat{\pi}$ on $\mathcal{M} \otimes \mathcal{N}$, and the conclusion of the corollary is obtained by setting $\mu \times \nu = \pi = \hat{\pi}$. ■

Theorem 3.10 (General Product Measure Theorem). *Let $(X_1, \mathcal{M}_1, \mu)$ be a σ-finite measure space, and let $(X_2, \mathcal{M}_2), \dots, (X_n, \mathcal{M}_n)$ be measurable spaces, and for each $1 \le k \le n-1$ let*

$$\mathcal{F}_k = \left\{ \mu(x_1, \dots, x_k, \cdot) : \mathcal{M}_{k+1} \to \overline{\mathbb{R}} \;\middle|\; (x_1, \dots, x_k) \in X_1 \times \cdots \times X_k \right\}$$

be a family of set functions with the following properties:

P1. *$\mu(x_1, \dots, x_k, \cdot)$ is a measure on $\mathcal{M}_{k+1}$ for each $(x_1, \dots, x_k) \in X_1 \times \cdots \times X_k$.*

P2. *For each $B \in \mathcal{M}_{k+1}$ we have*

$$\mu(\cdot, B) : (X_1 \times \cdots \times X_k, \mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_k) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}})).$$

P3. *$\mathcal{F}_k$ is uniformly σ-finite: There exist sequences $(B_{k+1,n})_{n\in\mathbb{N}} \subseteq \mathcal{M}_{k+1}$ and $(c_{k+1,n})_{n\in\mathbb{N}} \subseteq \mathbb{R}_+$ such that $X_{k+1} = \bigcup_{n=1}^{\infty} B_{k+1,n}$ and $\mu(\cdot, B_{k+1,n}) \le c_{k+1,n}$ for each n.*

Then there is a unique measure π on $\bigotimes_{k=1}^n \mathcal{M}_k$ such that

$$\pi\left(\prod_{k=1}^n A_k\right) = \int_{A_1} \int_{A_2} \cdots \int_{A_{n-1}} \mu(x_1, \dots, x_{n-1}, A_n)\, \mu(x_1, \dots, x_{n-2}, dx_{n-1}) \cdots \mu(x_1, dx_2)\, \mu(dx_1)$$

for each $\prod_{k=1}^n A_k \in \mathcal{R}(\prod_{k=1}^n X_k)$, and moreover the measure space $(\prod_{k=1}^n X_k, \bigotimes_{k=1}^n \mathcal{M}_k, \pi)$ is σ-finite.

In the case when $n = 3$ the conclusion of the General Product Measure Theorem states that $\pi : \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3 \to \overline{\mathbb{R}}$ is the unique measure such that

$$\begin{aligned} \pi(A_1 \times A_2 \times A_3) &= \int_{A_1} \left(\int_{A_2} \mu(x_1, x_2, A_3)\, \mu(x_1, dx_2) \right) \mu(dx_1) \\ &:= \int_{A_1} \int_{A_2} \mu(x_1, x_2, A_3)\, \mu(x_1, dx_2)\, \mu(dx_1) \end{aligned}$$

for each $A_1 \times A_2 \times A_3 \in \mathcal{R}(X_1 \times X_2 \times X_3)$. By definition,

$$\int_{A_2} \mu(x_1, x_2, A_3)\, \mu(x_1, dx_2) := \int_{A_2} \mu(x_1, \cdot, A_3)\, d\mu(x_1, \cdot),$$

where $x_1 \in X_1$ is taken to be fixed in the integration process.

Corollary 3.11 (**General Classical Product Measure Theorem**). *If* $(X_k, \mathcal{M}_k, \mu_k)$ *is a* σ*-finite measure space for each* $1 \leq k \leq n$*, then there is a unique measure* $\mu_1 \times \cdots \times \mu_n$ *on* $\bigotimes_{k=1}^n \mathcal{M}_k$ *such that*

$$(\mu_1 \times \cdots \times \mu_n)(A_1 \times \cdots \times A_n) = \prod_{k=1}^n \mu_k(A_k)$$

for all $A_1 \times \cdots \times A_n \in \mathcal{R}(X_1 \times \cdots \times X_n)$.

Given measures $\mu_1, \ldots, \mu_n$ on a single σ-algebra $\mathcal{M}$, the symbol $\prod_{k=1}^n \mu_k$ we take to represent the usual product of functions: $\prod_{k=1}^n \mu_k = \mu_1 \cdots \mu_n$. In contrast, given σ-finite measure spaces $(X_k, \mathcal{M}_k, \mu_k)$ for $1 \leq k \leq n$, we define

$$\bigtimes_{k=1}^n \mu_k = \mu_1 \times \cdots \times \mu_n,$$

which is the product measure on $\bigotimes_{k=1}^n \mathcal{M}_k$ whose existence and uniqueness is assured by the General Classical Product Measure Theorem.

Proposition 3.12. *For any* $n \in \mathbb{N}$*, if* $\lambda_n : \mathcal{B}(\mathbb{R}^n) \to \overline{\mathbb{R}}$ *is the Lebesgue measure on* $\mathcal{B}(\mathbb{R}^n)$*, then* $\bigtimes_{k=1}^n \lambda_1 = \lambda_n$ *on* $\mathcal{B}(\mathbb{R}^n)$.

Proof. Let $\lambda = \lambda_1$. By Corollary 3.4, $\mathcal{B}(\mathbb{R}^n) = \bigotimes_{k=1}^n \mathcal{B}(\mathbb{R})$, so certainly λ_n and $\bigtimes_{k=1}^n \lambda$ have the same domain. Define $F : \mathbb{R}^n \to \mathbb{R}$ by $F(\mathbf{x}) = \prod_{k=1}^n [\mathbf{x}]_k$, which is a distribution function. By definition, $\lambda_n : \mathcal{B}(\mathbb{R}^n) \to \overline{\mathbb{R}}$ is the unique (by Theorem 1.55) Lebesgue-Stieltjes measure such that $\lambda_n(\mathbf{a}, \mathbf{b}] = F(\mathbf{a}, \mathbf{b}]$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ with $\mathbf{a} \leq \mathbf{b}$. Suppose $I \subseteq \mathbb{R}^n$ is a bounded interval, so $I = \prod_{k=1}^n I_k$ with $I_k \subseteq \mathbb{R}$ a bounded interval for each k. Since $I_k \in \mathcal{B}(\mathbb{R})$ for each k, by Corollary 3.11 and Example 1.56 we have

$$\left(\bigtimes_{k=1}^n \lambda \right)(I) = \prod_{k=1}^n \lambda(I_k) = \lambda_n(I) < \infty,$$

and so $\bigtimes_{k=1}^n \lambda$ is a Lebesgue-Stieltjes measure on $\mathbb{R}^n$. In particular

$$\left(\bigtimes_{k=1}^n \lambda \right)(\mathbf{a}, \mathbf{b}] = \lambda_n(\mathbf{a}, \mathbf{b}] = F(\mathbf{a}, \mathbf{b}],$$

for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ with $\mathbf{a} \leq \mathbf{b}$, and so by the uniqueness provision of Theorem 1.55 we conclude that $\bigtimes_{k=1}^n \lambda = \lambda_n$ on $\mathcal{B}(\mathbb{R}^n)$. ∎

3.3 – Fubini's Theorem

The hypotheses stipulated in Fubini's Theorem and the General Fubini Theorem are identical to those in the Product Measure Theorem and General Product Measure Theorem, respectively.

Theorem 3.13 (Fubini's Theorem). *Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space, $(Y, \mathcal{N})$ a measurable space, and*

$$\mathcal{F} = \left\{\nu(x, \cdot) : \mathcal{N} \to \overline{\mathbb{R}} \;\middle|\; x \in X\right\}$$

a family of set functions with the following properties:

P1. *$\nu(x, \cdot)$ is a measure on $\mathcal{N}$ for each $x \in X$.*
P2. *$\nu(\cdot, B) : X \to \overline{\mathbb{R}}$ is Borel measurable for each $B \in \mathcal{N}$.*
P3. *$\mathcal{F}$ is uniformly σ-finite.*

Given $f : (X \times Y, \mathcal{M} \otimes \mathcal{N}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, define

$$F(x) = \int_Y f(x, y)\, \nu(x, dy).$$

Let $\mu \times \nu$ denote the unique measure π defined in Theorem 3.8.

1. *If $f \geq 0$, then $F : X \to \overline{\mathbb{R}}$ is an $(\mathcal{M}, \mathcal{B}(\overline{\mathbb{R}}))$-measurable function, and also*
$$\int_{X \times Y} f\, d(\mu \times \nu) = \int_X F\, d\mu = \int_X \left(\int_Y f(x, y)\, \nu(x, dy)\right) \mu(dx).$$
2. *Suppose $\int_{X \times Y} f\, d(\mu \times \nu)$ exists in $\overline{\mathbb{R}}$ (resp. $\mathbb{R}$), S is the set of all $x \in X$ for which $F(x)$ does not exist in $\overline{\mathbb{R}}$ (resp. $\mathbb{R}$), and $G : X \to \overline{\mathbb{R}}$ is any $(\mathcal{M}, \mathcal{B}(\overline{\mathbb{R}}))$-measurable function. Then $\mu(S) = 0$, the function $\hat{F} : X \to \overline{\mathbb{R}}$ given by*
$$\hat{F}(x) = \begin{cases} F(x), & x \in X \setminus S \\ G(x), & x \in S \end{cases}$$
is $(\mathcal{M}, \mathcal{B}(\overline{\mathbb{R}}))$-measurable, and
$$\int_{X \times Y} f\, d(\mu \times \nu) = \int_X \hat{F}\, d\mu. \tag{3.3}$$

In the second part of the Fubini'sTheorem it is common to set $G \equiv 0$ and write equation (3.3) as

$$\int_{X \times Y} f\, d(\mu \times \nu) = \int_X \left(\int_Y f(x, y)\, \nu(x, dy)\right) \mu(dx),$$

with the understanding that the value of the inner integral is to be taken to be 0 at all $x \in X$ where the integral fails to return a value in $\overline{\mathbb{R}}$ (resp. $\mathbb{R}$).

Corollary 3.14 (Tonelli's Theorem). *Suppose the hypotheses of the Fubini's Theorem hold. If*

$$\int_X \left(\int_Y |f(x, y)|\, \nu(x, dy)\right) \mu(dx) \in \mathbb{R},$$

then $\int_{X \times Y} f\, d(\mu \times \nu) \in \mathbb{R}$ as well.

Corollary 3.15 (**Classical Fubini Theorem**). *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ-finite measure spaces. If $f : (X \times Y, \mathcal{M} \otimes \mathcal{N}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, then*

$$\int_{X\times Y} f\, d(\mu \times \nu) = \int_X \left(\int_Y f(x,y)\, \nu(dy) \right) \mu(dx) = \int_Y \left(\int_X f(x,y)\, \mu(dx) \right) \nu(dy) \tag{3.4}$$

provided the integral at left exists.

Equation (3.4) may be written more simply as

$$\int_{X\times Y} f\, d(\mu \times \nu) = \int_X \int_Y f\, d\nu\, d\mu = \int_Y \int_X f\, d\mu\, d\nu.$$

Theorem 3.16 (**General Fubini Theorem**). *Let $(X_1, \mathcal{M}_1, \mu)$ be a σ-finite measure space, and let $(X_2, \mathcal{M}_2), \dots, (X_n, \mathcal{M}_n)$ be measurable spaces, and for each $1 \le k \le n-1$ let*

$$\mathcal{F}_k = \left\{ \mu(x_1, \dots, x_k, \cdot) : \mathcal{M}_{k+1} \to \overline{\mathbb{R}} \;\middle|\; (x_1, \dots, x_k) \in X_1 \times \cdots \times X_k \right\}$$

be a family of set functions with the following properties:

P1. *$\mu(x_1, \dots, x_k, \cdot)$ is a measure on $\mathcal{M}_{k+1}$ for each $(x_1, \dots, x_k) \in X_1 \times \cdots \times X_k$.*
P2. *For each $B \in \mathcal{M}_{k+1}$ we have*

$$\mu(\cdot, B) : (X_1 \times \cdots \times X_k, \mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_k) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}})).$$

P3. *$\mathcal{F}_k$ is uniformly σ-finite.*

Given $f : (\prod_{k=1}^n X_k, \bigotimes_{k=1}^n \mathcal{M}_k) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, define

$$F_k(x_1, \dots, x_k) = \int_{X_{k+1}} \cdots \int_{X_n} f(x_1, \dots, x_n)\, \mu(x_1, \dots, x_{n-1}, dx_n) \cdots \mu(x_1, \dots, x_k, dx_{k+1})$$

for each $1 \le k \le n-1$. Finally, for each $1 \le k \le n$ let π_k denote the unique measure on $\bigotimes_{j=1}^k \mathcal{M}_j$ considered in Theorem 3.10.

1. *If $f \ge 0$, then each $F_k : \prod_{j=1}^k X_j \to \overline{\mathbb{R}}$ is a $(\bigotimes_{j=1}^k \mathcal{M}_j, \mathcal{B}(\overline{\mathbb{R}}))$-measurable function, and also*

$$\int_{\prod_{k=1}^n X_k} f\, d\pi_n = \int_{X_1} \int_{X_2} \cdots \int_{X_n} f(x_1, \dots, x_n)\, \mu(x_1, \dots, x_{n-1}, dx_n) \cdots \mu(x_1, dx_2)\, \mu(dx_1). \tag{3.5}$$

2. *Suppose the integral at left in (3.5) exists in $\overline{\mathbb{R}}$ (resp. $\mathbb{R}$), S_k is the set of all $\mathbf{x} \in \prod_{j=1}^k X_j$ for which $F_k(\mathbf{x})$ does not exist in $\overline{\mathbb{R}}$ (resp. $\mathbb{R}$), and $G_k : \prod_{j=1}^k X_j \to \overline{\mathbb{R}}$ is any $(\bigotimes_{j=1}^k \mathcal{M}_j, \mathcal{B}(\overline{\mathbb{R}}))$-measurable function. Then $\pi_k(S_k) = 0$, the function $\hat{F}_k : \prod_{j=1}^k X_j \to \overline{\mathbb{R}}$ given by*

$$\hat{F}_k(\mathbf{x}) = \begin{cases} F_k(\mathbf{x}), & \mathbf{x} \in \left(\prod_{j=1}^k X_j \right) \setminus S_k \\ G(\mathbf{x}), & \mathbf{x} \in S_k \end{cases}$$

is $(\bigotimes_{j=1}^k \mathcal{M}_j, \mathcal{B}(\overline{\mathbb{R}}))$-measurable, and (3.5) still holds with the understanding that the F_k functions are replaced with the $\hat{F}_k$ functions.

Corollary 3.17 (**General Classical Fubini Theorem**). *For each $1 \leq k \leq n$ let $(X_k, \mathcal{M}_k, \mu_k)$ be a σ-finite measure space, and let $\mu_1 \times \cdots \times \mu_n$ denote the measure π_n in Theorem 3.16. If $f : (\prod_{k=1}^n X_k, \bigotimes_{k=1}^n \mathcal{M}_k) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, then*

$$\int_{\prod_{k=1}^n X_k} f \, d(\mu_1 \times \cdots \times \mu_n) = \int_{X_{\sigma(1)}} \cdots \int_{X_{\sigma(n)}} f \, d\mu_{\sigma(n)} \cdots d\mu_{\sigma(1)}$$

for any permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$, provided the integral at left exists.

Example 3.18. Show that if $f : [a, b] \times [c, d] \to \mathbb{R}$ is continuous, then

$$\int_a^b \int_c^d f(x, y) \, dy \, dx = \int_c^d \int_a^b f(x, y) \, dx \, dy. \tag{3.6}$$

Solution. Let $I = [a, b]$ and $J = [c, d]$, and recall the Lebesgue measure $\lambda : \mathcal{B}(\mathbb{R}) \to \overline{\mathbb{R}}$. The measure spaces $(I, \mathcal{B}(I), \lambda)$ and $(J, \mathcal{B}(J), \lambda)$ are both finite, and hence both σ-finite. By Proposition 2.4 the function f is $(\mathcal{B}(I \times J), \mathcal{B}(\mathbb{R}))$-measurable, and thus

$$f : (I \times J, \mathcal{B}(I) \otimes \mathcal{B}(J)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

since $\mathcal{B}(I \times J) = \mathcal{B}(I) \otimes \mathcal{B}(J)$. Indeed, by Proposition 2.2 we have

$$f : (I \times J, \mathcal{B}(I) \otimes \mathcal{B}(J)) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}})).$$

Now, the continuity of f on the compact set $I \times J \subseteq \mathbb{R}^2$ implies that f is bounded. Recalling the Lebesgue measure $\lambda_2 : \mathcal{B}(\mathbb{R}^2) \to \overline{\mathbb{R}}$, and noting that the measure space $(I \times J, \mathcal{B}(I \times J), \lambda_2)$ is finite, by Proposition 2.28 it follows that $\int_{I \times J} f \, d\lambda_2$ exists in $\mathbb{R}$. However,

$$\int_{I \times J} f \, d\lambda_2 = \int_{I \times J} f \, d(\lambda \times \lambda)$$

since $\lambda_2 = \lambda \times \lambda$ by Proposition 3.12, and so $\int_{I \times J} f \, d(\lambda \times \lambda)$ exists in $\mathbb{R}$. By the Classical Fubini Theorem, therefore,

$$\int_{I \times J} f \, d\lambda_2 = \int_I \left(\int_J f(x, y) \, d\lambda(y) \right) d\lambda(x) = \int_J \left(\int_I f(x, y) \, d\lambda(x) \right) d\lambda(y). \tag{3.7}$$

Define $\varphi : I \to \mathbb{R}$ and $\psi : J \to \mathbb{R}$ by

$$\varphi(x) = \int_J f(x, y) \, d\lambda(y) \quad \text{and} \quad \psi(y) = \int_I f(x, y) \, d\lambda(x).$$

Fix $x_0 \in I$. Let $\epsilon > 0$. The function f is in fact uniformly continuous on $I \times J$, so there exists some $\delta > 0$ such that, for all $(x_1, y_1), (x_2, y_2) \in I \times J$,

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} < \delta \;\;\Rightarrow\;\; |f(x_1, y_1) - f(x_2, y_2)| < \frac{\epsilon}{d - c}.$$

Let $x \in I$ be such that $|x - x_0| < \delta$. For any $y \in J$ the distance between (x, y) and (x_0, y) is less than δ, and so

$$\begin{aligned} |\varphi(x) - \varphi(x_0)| &= \left| \int_J f(x, y) \, d\lambda(y) - \int_J f(x_0, y) \, d\lambda(y) \right| \\ &= \left| \int_J \left[f(x, y) - f(x_0, y) \right] d\lambda(y) \right| && \text{Additivity Theorem} \end{aligned}$$

$$\leq \int_J \big|f(x,y) - f(x_0,y)\big|\, d\lambda(y) \qquad \text{Theorem 2.27(3)}$$

$$\leq \int_J \frac{\epsilon}{d-c}\, d\lambda(y) = \frac{\epsilon\lambda(J)}{d-c} = \epsilon, \qquad \text{Theorem 2.27(2)}$$

noting that $\lambda(J) = d - c$. This shows φ to be continuous at x_0, and hence φ is continuous on I. A similar argument will show that ψ is continuous on J. Now, by Theorem 2.52,

$$\int_I \left(\int_J f(x,y)\, d\lambda(y)\right) d\lambda(x) = \int_a^b \left(\int_J f(x,y)\, d\lambda(y)\right) dx = \int_a^b \left(\int_c^d f(x,y)\, dy\right) dx$$

and

$$\int_J \left(\int_I f(x,y)\, d\lambda(x)\right) d\lambda(y) = \int_c^d \left(\int_I f(x,y)\, d\lambda(x)\right) dy = \int_c^d \left(\int_a^b f(x,y)\, dx\right) dy$$

These results, together with (3.7), give (3.6) at last. ■