Chapter 4

Integration

4.1 Simple functions

Often, it is necessary to use the symbols $-\infty$ or ∞ in calculations. In these cases we work with the extended real line, i.e., $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{\infty\} = [-\infty, \infty]$. Functions that take values in $\bar{\mathbb{R}}$ are called *numerical* functions. The Borel sets associated with the extended real line are denoted by $\bar{\mathcal{B}} := \mathcal{B}(\bar{\mathbb{R}})$ and are defined as the collection of sets $\bar{B}$ such that $\bar{B} = B \cup S$ where $B \in \mathcal{B}(\mathbb{R})$ and $S \in \{\emptyset, \{-\infty\}, \{\infty\}, \{-\infty, \infty\}\}$. It can be verified that $\bar{\mathcal{B}}$ is a σ-algebra and that $\mathcal{B}(\mathbb{R}) = \mathbb{R} \cap \mathcal{B}(\bar{\mathbb{R}}) := \{\mathbb{R} \cup B : B \in \mathcal{B}(\mathbb{R})\}$. In addition, $\bar{\mathcal{B}}$ is generated by a collection of sets of the form $[a, \infty]$ (or $(a, \infty]$, $[-\infty, a]$, $[-\infty, a)$) where $a \in \mathbb{R}$.

Theorem 4.1. *$\bar{\mathcal{B}} = \sigma(\mathcal{C})$, where $\mathcal{C} := \{[a, \infty] : a \in \mathbb{R}\}$.*

Proof. Let $\mathcal{C} := \{[a, \infty] : a \in \mathbb{R}\}$ and $\mathcal{G} := \sigma(\mathcal{C})$. Note that since $[a, \infty] = [a, \infty) \cup \{\infty\}$, $[a, \infty] \in \bar{\mathcal{B}}$ and $\mathcal{C} \subset \bar{\mathcal{B}}$. Then, since $\bar{\mathcal{B}}$ is a σ-algebra $\sigma(\mathcal{C}) := \mathcal{G} \subset \bar{\mathcal{B}}$. Now, let $\mathcal{C}_1 = \{[a, b) : -\infty < a \leq b < \infty\}$ and note that $[a, b) = [a, \infty] - [b, \infty] \in \mathcal{G}$. Hence, $\mathcal{C}_1 \subset \mathcal{G}$ and $\sigma(\mathcal{C}_1) = \mathcal{B}(\mathbb{R}) \subset \mathcal{G}$ since $\mathcal{G}$ is a σ-algebra.

Note that $\{\infty\} = \bigcap\limits_{n \in \mathbb{N}} [n, \infty]$, $\{-\infty\} = \bigcap\limits_{n \in \mathbb{N}} [-\infty, -n) = \bigcap\limits_{n \in \mathbb{N}} [-n, \infty]^c$ and, consequently, $\{\infty\}, \{-\infty\} \in \mathcal{G}$. Then, for all $B \in \mathcal{B}(\mathbb{R})$ and $S \in \{\emptyset, \{-\infty\}, \{\infty\}, \{-\infty, \infty\}\}$ we have $B \cup S \in \mathcal{G}$, showing that $\bar{\mathcal{B}} \subset \mathcal{G}$. ■

Let $(\mathbb{X}, \mathcal{F})$ and $(\mathbb{R}, \mathcal{B})$ be measurable spaces. Since the indicator function of a measurable set is a measurable function, it follows from Theorem 3.5 that if $\{A_j\}_{j=1}^n$ with $n \in \mathbb{N}$ is a pairwise disjoint collection in $\mathcal{F}$ and $a_j \in \mathbb{R}$ for $j = 1, \cdots, n$, the linear combination

$$f(x) = \sum_{j=1}^{n} a_j I_{A_j}(x) \tag{4.1}$$

is a $\mathcal{F} - \mathcal{B}$-measurable function.

Definition 4.1. *A real-valued function on a measurable space $(\mathbb{X}, \mathcal{F})$ is said to be simple if it has the representation (4.1). A standard representation of a simple function is given by*

$$f(x) = \sum_{j=0}^{n} a_j I_{A_j}(x) \text{ with } a_0 = 0 \text{ and } A_0 = (\cup_{j=1}^n A_j)^c. \tag{4.2}$$

Remark 4.1.

1. *If $f : (\mathbb{X}, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$ is measurable and takes on finitely many values, say $\{a_j\}_{j=1}^n$ then it is a simple function. To see this, note that $B_j = \{x : f(x) = a_j\}$ is measurable, since $B_j = \{x : f(x) \leq a_j\} - \{x : f(x) < a_j\}$ and f is measurable. Also, note that the collection $\{B_j\}_{j=1}^n$ is pairwise disjoint. Hence,*

$$f(x) = \sum_{j=1}^{n} a_j I_{B_j}(x) = \sum_{j=1}^{n} a_j I_{\{x: f(x) = a_j\}}(x). \tag{4.3}$$

   Conversely, if f is simple it takes on finitely many values.

2. *Representation (4.2) is not unique, but a simple function has at least one representation such as (4.2) .*

The next theorem shows that certain functions of simple functions are simple functions.

Theorem 4.2. *Let $f : (\mathbb{X}, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$ and $g : (\mathbb{X}, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$ be simple functions. Then, $f \pm g$, cf for $c > 0$, fg, $f^+ = \max\{f, 0\}$, $f^- = -\min\{f, 0\}$ and $|f|$ are simple functions.*

4.2 Integral of simple functions

Definition 4.2. *Let $f : (\mathbb{X}, \mathcal{F}, \mu) \to (\mathbb{R}, \mathcal{B})$ be a non-negative simple function with standard representation (4.2). The integral of f with respect to μ, denoted by $\int_{\mathbb{X}} f d\mu$, is given by*

$$\int_{\mathbb{X}} f d\mu := \sum_{j=0}^{n} a_j \mu(A_j) \in [0, \infty]. \tag{4.4}$$

By definition $a_j \in \mathbb{R}$ for $j = 0, 1, \cdots, n$, but since μ takes values in $[0, \infty]$ we can have $\int_{\mathbb{X}} f d\mu = \infty$. If μ is a finite measure, e.g., a probability measure P, then it must be that $\int_{\mathbb{X}} f d\mu \in \mathbb{R}$. When $\mathbb{X} := \Omega$ an outcome space, $f := X$ is a random variable and $\mu := P$ is a probability measure we write $E_P(X) := \int_{\Omega} X dP$ and call it the expectation of X given probability P.

It will be convenient, in the case of simple functions, to write $I_\mu(f) := \int_{\mathbb{X}} f d\mu$.

Remark 4.2. *Since the representation (4.2) is not unique, for uniqueness, the definition of integral requires that it be invariant to the representation used. To see this, suppose that $f(x) = \sum_{j=0}^{n} a_j I_{A_j}(x) = \sum_{k=0}^{m} b_k I_{B_k}(x)$. Then, $\mathbb{X} = \cup_{j=0}^{n} A_j = \cup_{k=0}^{m} B_k$ and*

$$A_j = \cup_{k=0}^{m} (A_j \cap B_k),\ B_k = \cup_{j=0}^{n} (A_j \cap B_k).$$

Since μ is finitely additive and the sets in the above unions are disjoint we have that

$$\sum_{j=0}^{n} a_j \mu(A_j) = \sum_{j=0}^{n} a_j \sum_{k=0}^{m} \mu(A_j \cap B_k) = \sum_{j=0}^{n} \sum_{k=0}^{m} a_j \mu(A_j \cap B_k).$$

Similarly,

$$\sum_{k=0}^{m} b_k \mu(B_k) = \sum_{k=0}^{m} b_k \sum_{j=0}^{n} \mu(A_j \cap B_k) = \sum_{j=0}^{n} \sum_{k=0}^{m} b_k \mu(A_j \cap B_k).$$

But $a_j = b_k$ whenever $A_j \cap B_k \neq \emptyset$, and when $A_j \cap B_k = \emptyset$, $\mu(A_j \cap B_k) = 0$. Thus, $a_j \mu(A_j \cap B_k) = b_k \mu(A_j \cap B_k)$ for all pairs (j, k), and $I_\mu(f)$ is invariant to the representation of the simple function.

Theorem 4.3. *Let $f : (\mathbb{X}, \mathcal{F}, \mu) \to (\mathbb{R}, \mathcal{B})$ and $g : (\mathbb{X}, \mathcal{F}, \mu) \to (\mathbb{R}, \mathcal{B})$ be simple non-negative functions. Then,*

1. $\int_{\mathbb{X}} cf d\mu = c \int_{\mathbb{X}} f d\mu$ *for* $c \geq 0$ *and* $\int_{\mathbb{X}} I_E d\mu = \mu(E)$ *for* $E \in \mathcal{F}$.
2. $\int_{\mathbb{X}} (f+g) d\mu = \int_{\mathbb{X}} f d\mu + \int_{\mathbb{X}} g d\mu$,
3. *If for* $E \in \mathcal{F}$, *we define* $m(E) = \int_{\mathbb{X}} f I_E d\mu$, *then* m *is a measure on* $\mathcal{F}$.
4. $f \leq g \implies \int_{\mathbb{X}} f d\mu \leq \int_{\mathbb{X}} g d\mu$.

Proof. For 1., note that $c \geq 0 \implies cf \geq 0$ with representation $cf(x) = \sum_{j=0}^{n} c a_j I_{A_j}(x)$. Therefore, $\int_{\mathbb{X}} cf d\mu = \sum_{j=0}^{n} c a_j \mu(A_j) = c \sum_{j=0}^{n} a_j \mu(A_j) = c \int_{\mathbb{X}} f d\mu$. For the second part, note that $I_E(x) = I_E(x) + 0 I_{E^c}(x)$. Hence, $\int_{\mathbb{X}} I_E d\mu = \mu(E)$.

For 2., let $f(x) = \sum_{j=0}^{n} a_j I_{A_j}(x)$ and $g(x) = \sum_{k=0}^{m} b_k I_{B_k}(x)$. Then, $f(x) + g(x) = \sum_{j=0}^{n} \sum_{k=0}^{m} (a_j + b_k) I_{A_j \cap B_k}(x)$ with $(A_j \cap B_k) \cap (A_{j'} \cap B_{k'}) = \emptyset$ whenever $(j,k) \neq (j',k')$. Then,

$$
\begin{aligned}
\int_{\mathbb{X}} (f+g)\, d\mu &= \sum_{j=0}^{n} \sum_{k=0}^{m} (a_j + b_k) \mu(A_j \cap B_k) \\
&= \sum_{j=0}^{n} a_j \sum_{k=0}^{m} \mu(A_j \cap B_k) + \sum_{k=0}^{m} b_k \sum_{j=0}^{n} \mu(A_j \cap B_k) \\
&= \sum_{j=0}^{n} a_j \mu(A_j) + \sum_{k=0}^{m} b_k \mu(B_k),
\end{aligned}
$$

since $\mathbb{X}$ is the union of both $\{A_j\}$ and $\{B_k\}$. Then, by definition $\int_{\mathbb{X}} (f+g) d\mu = \int_{\mathbb{X}} f d\mu + \int_{\mathbb{X}} g d\mu$.

For 3., note that $f(x) I_E(x) = \sum_{j=0}^{n} a_j I_{A_j \cap E}(x)$. From parts 2. and 1.,

$$m(E) = \int_{\mathbb{X}} f I_E d\mu = \sum_{j=0}^{n} a_j \int_{\Omega} I_{A_j \cap E}(x) d\mu = \sum_{j=0}^{n} a_j \mu(A_j \cap E).$$

But $\mu(A_j \cap E)$ is a measure, and we have expressed $m(E)$ as a linear combination of measures on $\mathcal{F}$, hence m is a measure on $\mathcal{F}$.

For 4., write $g = f + (g - f)$. Note that $g - f$ is simple and non-negative since $g \geq f$. Hence, $I_\mu(g) = I_\mu(f) + I_\mu(g - f) \geq I_\mu(f)$. ■

4.3 Integral of non-negative functions

We start with the following fundamental theorem.

Theorem 4.4. *Let $f : (\Omega, \mathcal{F}) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ be a non-negative measurable function. Then, there exists a sequence $\varphi_n : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$ of simple non-negative functions such that:*

1. *$\varphi_n(\omega) \leq \varphi_{n+1}(\omega)$, for all $\omega \in \Omega$ and $n \in \mathbb{N}$*

2. *$\lim_{n\to\infty} \varphi_n(\omega) = f(\omega)$, for all $\omega \in \Omega$.*

Proof. 1. For each $n = 1, 2, \cdots$ define the sets

$$E_{k,n} = \begin{cases} \{\omega \in \Omega : \frac{k}{2^n} \leq f(\omega) < \frac{k}{2^n} + \frac{1}{2^n}\} = f^{-1}([\frac{k}{2^n}, \frac{k}{2^n} + \frac{1}{2^n})) \text{ for } k = 0, 1, \cdots, n2^n - 1 \\ \{\omega \in \Omega : f(\omega) \geq n\} = f^{-1}([n, \infty]) \text{ for } k = n2^n. \end{cases}$$

For each n, the sets $\{E_{k,n} : k = 0, 1, \cdots, n2^n\}$ are disjoint by construction, belong to $\mathcal{F}$ since f is measurable and $\cup_{k=0}^{n2^n} E_{k,n} = \Omega$. Now, let

$$\varphi_n(\omega) = \sum_{k=0}^{n2^n} \frac{k}{2^n} I_{E_{k,n}}(\omega).$$

Fix $\omega \in \Omega$ and for any $n \in \mathbb{N}$ we note that $\omega \in E_{k_0,n}$ for some k_0. By definition

$$\varphi_n(\omega) = \begin{cases} \frac{k_0}{2^n} \text{ if } k_0 = 0, 1, \cdots, n2^n - 1 \\ n \text{ if } k_0 = n2^n. \end{cases}$$

First, let $k_0 \in \{0, 1, \cdots, n2^n - 1\}$ and consider $n+1$. The lower bound on $[\frac{k_0}{2^n}, \frac{k_0}{2^n} + \frac{1}{2^n})$ must coincide with $\frac{k}{2^{n+1}}$, which gives $k = 2k_0$. Thus, $E_{k,n+1} = E_{2k_0,n+1} = f^{-1}\left([\frac{2k_0}{2^{n+1}}, \frac{2k_0}{2^{n+1}} + \frac{1}{2^{n+1}})\right) = f^{-1}\left([\frac{k_0}{2^n}, \frac{k_0}{2^n} + \frac{1}{2^{n+1}})\right)$ and

$$E_{k+1,n+1} = E_{2k_0+1,n+1} = f^{-1}\left([\frac{k_0}{2^n} + \frac{1}{2^{n+1}}, \frac{k_0}{2^n} + \frac{2}{2^{n+1}})\right) = f^{-1}\left([\frac{k_0}{2^n} + \frac{1}{2^{n+1}}, \frac{k_0}{2^n} + \frac{1}{2^n})\right).$$

Consequently, $E_{k_0,n} = E_{k,n+1} \cup E_{k+1,n+1} = E_{2k_0,n+1} \cup E_{2k_0+1,n+1}$. If $\omega \in E_{2k_0,n+1} \subset E_{k_0,n}$ then $\varphi_{n+1}(\omega) = \frac{2k_0}{2^{n+1}}$ and $\varphi_{n+1}(\omega) - \varphi_n(\omega) = \frac{2k_0}{2^{n+1}} - \frac{k_0}{2^n} = 0$. Alternatively, if $\omega \in E_{2k_0+1,n+1}$ then $\varphi_{n+1}(\omega) = \frac{2k_0+1}{2^{n+1}}$ and $\varphi_{n+1}(\omega) - \varphi_n(\omega) = \frac{2k_0+1}{2^{n+1}} - \frac{k_0}{2^n} = \frac{1}{2^{n+1}} > 0$. Consequently, if $\omega \in E_{k_0,n}$ then $\varphi_{n+1}(\omega) - \varphi_n(\omega) \geq 0$.

Second, if $k_0 = n2^n$ then $E_{k_0,n} = f^{-1}([n,\infty])$. Now, if $\omega \in f^{-1}([n+1,\infty])$ then $\varphi_{n+1}(\omega) = n+1$ and $\varphi_n(\omega) = n$. Consequently, $\varphi_{n+1}(\omega) - \varphi_n(\omega) = 1 > 0$. If $\omega \in f^{-1}([n, n+1])$ then $\varphi_n(\omega) = n$ and $\varphi_{n+1}(\omega) = \frac{k}{2^{n+1}}$ if $\omega \in f^{-1}([\frac{k}{2^{n+1}}, \frac{k}{2^{n+1}} + \frac{1}{2^{n+1}}))$. Setting the lower bound of the interval equal to n gives $k = n2^{n+1}$ and $\varphi_{n+1}(\omega) = n$ if $\omega \in f^{-1}([n, n+\frac{1}{2^{n+1}}))$, giving $\varphi_{n+1}(\omega) - \varphi_n(\omega) = 0$. If $\omega \in f^{-1}([n+\frac{1}{2^{n+1}}, n+\frac{2}{2^{n+1}}))$ then $\varphi_{n+1}(\omega) = \frac{n2^{n+1}+1}{2^{n+1}}$ and consequently $\varphi_{n+1}(\omega) - \varphi_n(\omega) = \frac{1}{2^{n+1}} > 0$. Continuing in this fashion for subsequent sub-intervals of $[n, n+1]$ gives $\varphi_{n+1}(\omega) - \varphi_n(\omega) \geq 0$.

2. From item 1, we have that $\varphi_1(\omega) \leq \varphi_2(\omega) \leq \cdots \leq f(\omega)$ for all $\omega \in \Omega$. Hence, $\lim_{n\to\infty} \varphi_n(\omega) = \sup_{n\in\mathbb{N}} \varphi_n(\omega)$. But $0 \leq f(\omega) - \varphi_n(\omega) \leq \frac{1}{2^n}$ and taking limits as $n \to \infty$ we have $f(\omega) = \lim_{n\to\infty} \varphi_n(\omega) = \sup_{n\in\mathbb{N}} \varphi_n(\omega)$. ■

Definition 4.3. *Let $f : (\mathbb{X}, \mathcal{F}, \mu) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ be a non-negative measurable function. The integral of f with respect to μ is given by*

$$\int_{\mathbb{X}} f d\mu := \sup_{\varphi} \int_{\mathbb{X}} \varphi(x) d\mu := \sup_{\varphi} I_\mu(\varphi) \; \in [0,\infty], \tag{4.5}$$

where the sup *is taken over all simple functions φ which are non-negative satisfying $\varphi(x) \leq f(x)$ for all $x \in \mathbb{X}$.*

Remark 4.3. *If f is a non-negative simple function $\int_{\mathbb{X}} f d\mu = I_\mu(f)$.*

Theorem 4.5. *(Beppo-Levi Theorem) Let $(\mathbb{X}, \mathcal{F}, \mu)$ be a measure space and $\{f_j\}_{j\in\mathbb{N}}$ be an increasing sequence of non-negative measurable functions $f_j : (\mathbb{X}, \mathcal{F}) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$. Then $f = \sup_{j\in\mathbb{N}} f_j$ is a non-negative measurable function and*

$$\int_{\mathbb{X}} f d\mu := \int_{\mathbb{X}} \sup_{j\in\mathbb{N}} f_j d\mu = \sup_{j\in\mathbb{N}} \int_{\mathbb{X}} f_j d\mu.$$

Proof. That f is a non-negative measurable function follows from Theorem 3.6. Note that if g and h are non-negative measurable functions, we have by definition that

$$\int_{\mathbb{X}} g d\mu := \sup_{\varphi} \int_{\mathbb{X}} \varphi d\mu \text{ where } \varphi \leq g,\ \varphi \text{ a simple function.}$$

But if $g \leq h$,

$$\int_{\mathbb{X}} g d\mu \leq \sup_{\varphi} \int_{\mathbb{X}} \varphi d\mu = \int_{\mathbb{X}} h d\mu \text{ where } \varphi \leq h.$$

Now, $f_j \leq f := \sup_{j \in \mathbb{N}} f_j$. By the monotonicity of integrals, which we just established,

$$\int_{\mathbb{X}} f_j d\mu \leq \int_{\mathbb{X}} f d\mu.$$

Taking $\sup_{j \in \mathbb{N}}$ on both sides gives $\sup_{j \in \mathbb{N}} \int_{\mathbb{X}} f_j d\mu \leq \int_{\mathbb{X}} f d\mu$.

Now, we establish the reverse inequality, i.e., $\sup_{j \in \mathbb{N}} \int_{\mathbb{X}} f_j d\mu \geq \int_{\mathbb{X}} f d\mu$. Let $\varphi(x)$ be a simple non-negative function such that $\varphi \leq f$. If we can show that

$$I_\mu(\varphi) = \int_{\mathbb{X}} \varphi d\mu \leq \sup_{j \in \mathbb{N}} \int_{\mathbb{X}} f_j d\mu \tag{4.6}$$

we will have the desired inequality since we can take sup over all simple functions on both sides of (4.6) to give

$$\sup_{\varphi} \int_{\mathbb{X}} \varphi d\mu := \int_{\mathbb{X}} f d\mu \ \leq \sup_{j \in \mathbb{N}} \int_{\mathbb{X}} f_j d\mu.$$

Let φ be a simple non-negative function such that $\varphi \leq f$. Since $f(x) := \sup_{j \in \mathbb{N}} f_j(x)$, for every $x \in \mathbb{X}$ and $\epsilon \in (0, 1)$, there exists $N_{(x,\epsilon)}$ such that

$$f_j(x) \geq \epsilon\varphi(x) \text{ whenever } j \geq N_{(x,\epsilon)}.$$

Now, if $A_j = \{x : f_j(x) \geq \epsilon\varphi(x)\}$ we note that the sets A_j increase as $j \to \infty$ since $f_1 \leq f_2 \cdots$. Furthermore, these sets are measurable by measurability of f_j and φ. By definition of A_j

$$\epsilon I_{A_j}(x)\varphi(x) \leq I_{A_j}(x) f_j(x) \leq f_j(x). \tag{4.7}$$

Since φ is a simple function it has a standard representation $\varphi(x) = \sum_{i=0}^{m} y_i I_{B_i}(x)$ and

$$\epsilon I_{A_j}(x) \sum_{i=0}^{m} y_i I_{B_i}(x) = \epsilon \sum_{i=0}^{m} y_i I_{B_i \cap A_j}(x).$$

Thus, the integral of the simple function in this expression is given by $\epsilon \sum_{i=0}^{m} y_i \mu(B_i \cap A_j)$. By monotonicity of integrals and using (4.7) we have

$$\epsilon \sum_{i=0}^{m} y_i \mu(B_i \cap A_j) \leq \int_{\mathbb{X}} f_j d\mu \leq \sup_{j \in \mathbb{N}} \int_{\mathbb{X}} f_j d\mu.$$

Since $\varphi \leq f$, the collection $\{A_j\}$ grows to $\mathbb{X}$ as $j \to \infty$. Thus, by the fact that μ is continuous from below

$$\mu(B_i \cap A_j) \uparrow \mu(B_i \cap \mathbb{X}) = \mu(B_i) \text{ as } j \to \infty$$

and

$$\epsilon \sum_{i=0}^{m} y_i \mu(B_i) = \epsilon \int_{\mathbb{X}} \varphi d\mu \leq \sup_{j \in \mathbb{N}} \int_{\mathbb{X}} f_j d\mu.$$

Now, just let ϵ be arbitrarily close to 1 to finish the proof. ■

Remark 4.4. *1. If we take $f_j = \varphi_j$ where φ_j are non-negative simple functions and $f = \sup_{j \in \mathbb{N}} \varphi_j$, then*

$$\int_{\mathbb{X}} f d\mu = \sup_{j \in \mathbb{N}} \int_{\mathbb{X}} \varphi_j d\mu.$$

Note that sup *can be replaced with* $\lim_{j \to \infty}$.

2. If $E \in \mathcal{F}$, then $I_E(x) f(x)$ is a non-negative measurable function if $f \geq 0$. We define

$$\int_E f d\mu = \int_{\mathbb{X}} I_E f d\mu. \tag{4.8}$$

Theorem 4.6. *Let $(\mathbb{X}, \mathcal{F}, \mu)$ be a measure space and f, $g : (\mathbb{X}, \mathcal{F}, \mu) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ be numerical non-negative measurable functions. Then*

1. $\int_{\mathbb{X}} a f d\mu = a \int_{\mathbb{X}} f d\mu$ *for* $a \geq 0$,

2. $\int_{\mathbb{X}}(f+g)d\mu = \int_{\mathbb{X}} f d\mu + \int_{\mathbb{X}} g d\mu$,

3. *If* $E, F \in \mathcal{F}$ *and* $E \subset F$, *then* $\int_E f d\mu \le \int_F f d\mu$.

Proof. 1. If $a > 0$, let φ_n be an increasing sequence of measurable non-negative simple functions converging to f (such sequence exists by Theorem 4.4). Then, $a\varphi_n$ is an increasing sequence converging point wise to af. By Theorem 4.5 and the fact that $I_\mu(a\varphi_n) = aI_\mu(\varphi_n)$

$$\int_{\mathbb{X}} af d\mu = \lim_{n\to\infty}\int_{\mathbb{X}} a\varphi_n d\mu = a\lim_{n\to\infty}\int_{\mathbb{X}}\varphi_n(\omega)d\mu = a\int_{\mathbb{X}} f d\mu$$

2. Let φ_n, ψ_n be non-negative increasing simple functions converging to f and g. Then $\varphi_n + \psi_n$ is an increasing sequence converging to $f+g$. Again, by Theorem 4.5

$$\begin{aligned}\int_{\mathbb{X}}(f+g)d\mu &= \lim_{n\to\infty}\int_{\mathbb{X}}(\varphi_n+\psi_n)d\mu \text{ by Beppo-Levi's Theorem}\\ &= \lim_{n\to\infty}\int_{\mathbb{X}}\varphi_n d\mu + \lim_{n\to\infty}\int_{\mathbb{X}}\psi_n d\mu \text{ by Theorem 4.3}\\ &= \int_{\mathbb{X}} f d\mu + \int_{\mathbb{X}} g d\mu. \text{ by Beppo-Levi's Theorem}\end{aligned}$$

3. Since f is non-negative $fI_E \le fI_F$ therefore

$$\int_E f d\mu = \int_{\mathbb{X}} fI_E d\mu \le \int_{\mathbb{X}} fI_F d\mu = \int_F f d\mu.$$

■

Corollary 4.1. *Let* $\{f_j\}_{j\in\mathbb{N}}$ *be a sequence of measurable non-negative numerical functions, i.e.,* $f_j : (\mathbb{X}, \mathcal{F}, \mu) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$. *Then,* $\sum_{j=1}^{\infty} f_j$ *is measurable and*

$$\int_{\mathbb{X}}\left(\sum_{j=1}^{\infty} f_j\right)d\mu = \sum_{j=1}^{\infty}\int_{\mathbb{X}} f_j d\mu.$$

Proof. Let $S_m = \sum_{j=1}^{m} f_j$, $S = \lim_{m\to\infty}\sum_{j=1}^{m} f_j = \sum_{j=1}^{\infty} f_j$ and note that $0 \le S_1 \le S_2 \le \cdots$. Then, by Theorem 4.6.3 we have that

$$\int_{\mathbb{X}} S_m d\mu = \sum_{j=1}^{m}\int_{\mathbb{X}} f_j d\mu.$$

Taking limits as $m \to \infty$ and using Theorem 4.5, we have

$$\lim_{m\to\infty}\int_{\mathbb{X}} S_m d\mu = \lim_{m\to\infty}\sum_{j=1}^{m}\int_{\mathbb{X}} f_j d\mu = \sum_{j=1}^{\infty}\int_{\mathbb{X}} f_j d\mu = \int_{\mathbb{X}} S d\mu = \int_{\mathbb{X}}\left(\sum_{j=1}^{\infty} f_j\right) d\mu.$$

■

Theorem 4.7. *(Fatou's Lemma): Let* $\{f_j\}_{j\in\mathbb{N}}$ *be a sequence of measurable non-negative numerical functions* $f_j : (\mathbb{X}, \mathcal{F}, \mu) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$*. Then,* $f := \liminf_{j\to\infty} f_j$ *is measurable and*

$$\int_{\mathbb{X}} f d\mu \leq \liminf_{j\to\infty}\int_{\mathbb{X}} f_j d\mu.$$

Proof. First, f is measurable by Theorem 3.6. Let $g_n = \inf\{f_n, f_{n+1}, \cdots\}$ for $n = 1, 2, \cdots$, and note that $g_1 \leq f_1$, $g_1 \leq f_2, \cdots$. Also, $g_2 \leq f_2$, $g_2 \leq f_3 \cdots$. Thus, $g_n \leq f_j$ for all $n \leq j$. Furthermore, $g_1 \leq g_2 \leq \cdots$. Now, recall that $f := \liminf_{j\to\infty} f_j := \sup_{n\in\mathbb{N}} \inf_{j\geq n} f_j$ and

$$\lim_{n\to\infty} g_n = \liminf_{j\to\infty} f_j := f.$$

Also, $\int_{\mathbb{X}} g_n d\mu \leq \int_{\mathbb{X}} f_j d\mu$ for all $n \leq j$ and

$$\int_{\mathbb{X}} g_n d\mu \leq \liminf_{j\to\infty}\int_{\mathbb{X}} f_j d\mu.$$

Since the sequence $g_n \uparrow \liminf_{j\to\infty} f_j$, by Theorem 4.5

$$\lim_{n\to\infty}\int_{\mathbb{X}} g_n d\mu = \int_{\mathbb{X}} f d\mu \leq \liminf_{j\to\infty}\int_{\mathbb{X}} f_j(\omega) d\mu.$$

■

4.4 Integral of functions

Let $f : (\mathbb{X}, \mathcal{F}, \mu) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ be a measurable numerical function and $f^+ = \max\{f, 0\}$ and $f^- = -\min\{f, 0\}$.

Definition 4.4. *Let* $f : (\mathbb{X}, \mathcal{F}, \mu) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ *be a measurable numerical function such that* $\int_{\mathbb{X}} f^+ d\mu < \infty$ *and* $\int_{\mathbb{X}} f^- d\mu < \infty$. *In this case, we say that* f *is* μ*-integrable and we write*

$$\int_{\mathbb{X}} f d\mu := \int_{\mathbb{X}} f^+ d\mu - \int_{\mathbb{X}} f^- d\mu.$$

We note that $\int_{\mathbb{X}} f d\mu \in \mathbb{R}$ and denote by $\mathcal{L}_{\mathbb{R}}$ the set of integrable real functions and $\mathcal{L}_{\bar{\mathbb{R}}}$ the set of integrable numerical functions. A non-negative function f is said to be integrable if, and only if, $\int_{\mathbb{X}} f d\mu < \infty$. If $(\mathbb{X}, \mathcal{F}, \mu) := (\mathbb{R}^n, \mathcal{B}^n, \lambda^n)$ we call $\int_{\mathbb{R}^n} f d\lambda^n$ the Lebesgue integral.

Theorem 4.8. *Let* $f : (\mathbb{X}, \mathcal{F}, \mu) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ *be a measurable function. Then, the following statements are equivalent:*

1. $f \in \mathcal{L}_{\bar{\mathbb{R}}}$,

2. $|f| \in \mathcal{L}_{\bar{\mathbb{R}}}$,

3. *there exists* $0 \leq g \in \mathcal{L}_{\bar{\mathbb{R}}}$ *such that* $|f| \leq g$.

Proof. ($1 \implies 2$) Since, $|f| = f^+ + f^-$ and since integrability of f implies $\int_{\mathbb{X}} f^+ d\mu < \infty$ and $\int_{\mathbb{X}} f^- d\mu < \infty$ we have $\int_{\mathbb{X}} |f| d\mu = \int_{\mathbb{X}} f^+ d\mu + \int_{\mathbb{X}} f^- d\mu < \infty$.
($2 \implies 3$) Just take $g = |f|$.
($3 \implies 1$) Since $f^+ \leq |f| \leq g$ and $f^- \leq |f| \leq g$, we have by the monotonicity of the integral of non-negative functions and the integrability of g that $f^+, f^- \in \mathcal{L}_{\bar{\mathbb{R}}}$. Hence, $f \in \mathcal{L}_{\bar{\mathbb{R}}}$. ■

Theorem 4.9. *Let Let* $f : (\mathbb{X}, \mathcal{F}, \mu) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ *be a measurable function and assume that* μ *is a finite measure. Then,*

$$\int_{\mathbb{X}} |f| d\mu < \infty \iff \forall \epsilon > 0 \, \exists \, \delta > 0 \ni \int_{\mathbb{X}} |f| I_E d\mu < \epsilon, \forall E \ni \mu(E) < \delta.$$

Proof. ($\Longleftarrow$) Let $A_b = \{x : |f(x)| \leq b\}$ for $b > 0$. Since $\mathbb{X} = A_b \cup A_b^c$, choose b such that $\mu(A_b^c) < \delta$. Then, since $\mu(A_b^c) < \delta$ and μ is finite

$$\int_{\mathbb{X}} |f| d\mu = \int_{A_b} |f| d\mu + \int_{A_b^c} |f| d\mu \leq b\mu(A_b) + \epsilon < \infty.$$

($\Longrightarrow$) Since $\int_{\mathbb{X}} |f| d\mu < \infty$, for any $\epsilon > 0$, there exists $b > 0$ such that $\int_{\mathbb{X}} |f| I_{A_b^c} d\mu < \epsilon$. Now, for any measurable set E,

$$E = (A_b \cup A_b^c) \cap E = (A_b \cap E) \cup (A_b^c \cap E) \subset (A_b \cup E) \cup A_b^c.$$

Hence, $I_E \leq I_{(A_b \cup E) \cup A_b^c} = I_{(A_b \cup E)} + I_{A_b^c}$, where the equality follows from the fact that the two sets in the union are disjoint. Then,

$$\int_{\mathbb{X}} |f| I_E d\mu \leq \int_{\mathbb{X}} |f| I_{A_b \cup E} d\mu + \int_{\mathbb{X}} |f| I_{A_b^c} d\mu < b\mu(E) + \epsilon < 2\epsilon$$

where the last inequality follows if $\mu(E) < \delta = \frac{\epsilon}{b}$. ■

Theorem 4.10. *Let $f, g : (\mathbb{X}, \mathcal{F}, \mu) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ be measurable functions such that $f, g \in \mathcal{L}_{\bar{\mathbb{R}}}$ and $a \in \mathbb{R}$. Then,*

1. *$af \in \mathcal{L}_{\bar{\mathbb{R}}}$ and $\int_{\mathbb{X}} af d\mu = a \int_{\mathbb{X}} f d\mu$,*
2. *$(f + g) \in \mathcal{L}_{\bar{\mathbb{R}}}$ and $\int_{\mathbb{X}} (f + g) d\mu = \int_{\mathbb{X}} f d\mu + \int_{\mathbb{X}} g d\mu$,*
3. *$\max\{f, g\}$, $\min\{f, g\} \in \mathcal{L}_{\bar{\mathbb{R}}}$,*
4. *if $f \leq g$ then $\int_{\mathbb{X}} f d\mu \leq \int_{\mathbb{X}} g d\mu$.*

Proof. Homework. Use Theorems 4.8 and 4.6. ■

Remark 4.5. *Note that*

$$\left| \int_{\mathbb{X}} f d\mu \right| \leq \left| \int_{\mathbb{X}} f^+ d\mu \right| + \left| \int_{\mathbb{X}} f^- d\mu \right| = \int_{\mathbb{X}} f^+ d\mu + \int_{\mathbb{X}} f^- d\mu = \int_{\mathbb{X}} (f^+ + f^-) d\mu = \int_{\mathbb{X}} |f| d\mu.$$

Theorem 4.11. *Let $f : (\mathbb{X}, \mathcal{F}, \mu) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ be a non-negative measurable function such that $f \in \mathcal{L}_{\bar{\mathbb{R}}}$ and*

$$m(E) = \int_E f d\mu \text{ for all } E \in \mathcal{F}.$$

Then, m is a measure on $\mathcal{F}$.

Proof. Since $f \geq 0$, $m(E) \geq 0$. If $E = \emptyset$, then $fI_E = 0$ and

$$m(\emptyset) = \int_{\emptyset} f d\mu = \int_{\mathbb{X}} f I_{\emptyset} d\mu = \int_{\mathbb{X}} 0 d\mu = 0.$$

Now, let $\{E_j\}_{j\in\mathbb{N}}$ be a disjoint collection of sets in $\mathcal{F}$ such that $\cup_{j=1}^{\infty} E_j = E$ and let $f_n(x) = \sum_{j=1}^{n} f(x) I_{E_j}(x)$. By Theorem 4.6 $\int_{\mathbb{X}} f_n d\mu = \sum_{j=1}^{n} \int_{\mathbb{X}} f I_{E_j} d\mu$. Thus, $\int_{\mathbb{X}} f_n d\mu = \sum_{j=1}^{n} m(E_j)$. Note that $f_1 \leq f_2 \leq \cdots$ and converges to fI_E. Hence, by Theorem 4.5

$$m(E) = \int_{\mathbb{X}} f I_E d\mu = \lim_{n\to\infty} \int_{\mathbb{X}} f_n d\mu = \lim_{n\to\infty} \sum_{j=1}^{n} m(E_j) = \sum_{j=1}^{\infty} m(E_j).$$

■

Remark 4.6. *1. Suppose $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a random variable and P_X is the probability measure induced by X on $\mathcal{B}(\mathbb{R})$ as in Example 3.2. Then, in Theorem 4.11 letting $(\mathbb{X}, \mathcal{F}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$, we conclude that*

$$m_X(B) = \int_B f dP_X \text{ for all } B \in \mathcal{B}(\mathbb{R})$$

is a measure on $\mathcal{B}(\mathbb{R})$. In particular, if $B = (-\infty, x]$ for $x \in \mathbb{R}$, $m_X((-\infty, x]) = \int_{(-\infty,x]} f dP_X$.

2. m is called the measure with density function f with respect to μ and is denoted by $m = f\mu$. If m has a density with respect to μ it is traditional in mathematics to write $dm/d\mu$ for the the density function. We note that with a little more work we can recognize f as the Radon-Nikodým derivative of m with respect to the measure μ.

4.5 Exercises

1. Prove Theorem 4.2.

2. Show that if f is a non-negative measurable simple function, its integral, as defined in Definition 4.3 is equal to $I_\mu(f)$.

3. Let $(\mathbb{X}, \mathcal{F})$ be a measurable space and $\{\mu_n\}_{n\in\mathbb{N}}$ be a sequence of measures defined on it. Noting that $\mu = \sum_{n\in\mathbb{N}} \mu_n$ is also a measure on $(\mathbb{X}, \mathcal{F})$ (you don't have to prove this), show that

$$\int_{\mathbb{X}} f d\mu = \sum_{n\in\mathbb{N}} \int_{\mathbb{X}} f d\mu_n$$

for f non-negative and measurable.

4. Let $(\mathbb{X}, \mathcal{F}, \mu)$ be a measure space and $f : (\mathbb{X}, \mathcal{F}, \mu) \to (\mathbb{R}, \mathcal{B})$ be measurable and non-negative. For every $F \in \mathcal{F}$ consider $\int I_F f d\mu$. Is this a measure?

5. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{F}$.

   (a) Prove that $I_{\liminf_{n\to\infty} F_n} = \liminf_{n\to\infty} I_{F_n}$ and $I_{\limsup_{n\to\infty} F_n} = \limsup_{n\to\infty} I_{F_n}$.

   (b) Prove that $P\left(\liminf_{n\to\infty} F_n\right) \leq \liminf_{n\to\infty} P(F_n)$.

   (c) Prove that $\limsup_{n\to\infty} P(F_n) \leq P\left(\limsup_{n\to\infty} F_n\right)$.