Chapter 3 INTEGRATION THEORY

References

Textbooks:

1. G. B. Folland, *Real Analysis. Modern Techniques and Their Applications*, 2nd ed. John Wiley & Sons, 1999 (**pp. 43–69**)
2. H. L. Royden, P. M. Fitzpatrick, *Real Analysis*, 4th ed. Pearson Education, 2010 (**pp. 359–381**)
3. E. Kopp, J. Malczak, T. Zastawniak *Probability for Finance*, Cambridge University Press, 2014

By the end of the nineteenth century, some inadequacies in the Riemann theory of integration had become apparent.

- The collection of Riemann integrable functions became inconveniently small as mathematics developed.
- Limits of sequences of Riemann integrable functions are not necessarily Riemann integrable.

These inadequacies led others to invent other integration theories, the best known of which was due to Henri Lebesgue.

The Lebesgue theory of integration has become pre-eminent in contemporary mathematical research, since it enables one to integrate a much larger collection of functions, and to take limits of integrals more freely.

In this chapter, we develop the theory of integration on abstract measure spaces, paying particular attention to the Lebesgue integral on $\mathbb{R}$ and its generalization to $\mathbb{R}^n$.

In the study of metric spaces, continuous functions play an important role.

Analogously, in the study of measurable spaces, measurable functions are important.

Definition 1.1

Let $(X, \mathcal{M})$ be a measurable space. An extended real-valued function $f : X \to \overline{\mathbb{R}}$ is **measurable** (or **$\mathcal{M}$-measurable**) if

$$\{x \in X : f(x) < \alpha\} = f^{-1}([-\infty, \alpha))$$

is a measurable set for every real number α.

Example 1.1

(a) Every constant function on X is measurable.

(b) If $\mathcal{M} = \{\emptyset, X\}$, then only the constant functions are measurable.

(c) If $\mathcal{M}$ consists of all subsets of X, $\mathcal{M} = \mathcal{P}(X)$, then every function from X to $\overline{\mathbb{R}}$ is measurable.

Note If $(X, \mathcal{M}, \mu)$ is a probability space, a measurable function $\xi : X \to \mathbb{R}$ is also called a **random variable**.

Example 1.2 Let A be a subset of X. The function

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

is called the **indicator function** (sometimes also the **characteristic function**) of A.

- $\chi_A(x)$ is measurable $\iff$ A is measurable.
- If $(X, \mathcal{M}, \mu)$ is a probability space and $0 < \mu(A) < 1$, then χ_A is called a **Bernoulli random variable**.

Note $\boldsymbol{A \subset B \iff \chi_A \leq \chi_B}$

Sometimes we wish to consider measurability on subsets of X.

If f is a function on a measurable set A, we say that f is **measurable on *A*** if $\{x \in A : f(x) < \alpha\}$ is measurable for every real number α.

$$f \text{ is measurable on } A \in \mathcal{M}$$
$$\iff \{x \in A : f(x) < \alpha\} \in \mathcal{M} \quad \forall \alpha \in \mathbb{R}.$$

For instance,

If μ is complete and $\mu(A) = 0$, then every function defined on A is measurable on A.

Note

If f is measurable on X and A is a measurable subset of X, then f is measurable on A.

Theorem 1.1

Let A be a measurable set. For a function $f : A \to \overline{\mathbb{R}}$, the following statements are equivalent:

(a) *f is measurable.*

(b) *$\{x \in A : f(x) \leq \alpha\}$ is a measurable set for every real number α.*

(c) *$\{x \in A : f(x) > \alpha\}$ is a measurable set for every real number α.*

(d) *$\{x \in A : f(x) \geq \alpha\}$ is a measurable set for every real number α.*

If f is measurable, then $\{x \in A : f(x) = \alpha\}$ is measurable for each extended number α.

Example 1.3 Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\xi : \Omega \to \mathbb{R}$ is a measurable function.

ξ is said to be a **binomial random variable** with parameters $n \in \mathbb{N}$ and $p \in [0, 1]$ if

$$P\Big(\{\omega \in \Omega : \xi(x) = k\}\Big) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k},$$

$$k = 0, 1, 2, \ldots, n.$$

Remark 1.1

A finite function $f : A \to \mathbb{R}$ is measurable if and only if $f^{-1}(B) \in \mathcal{M}$ for every Borel set.

$$f : A \to \mathbb{R} \text{ is measurable } \iff f^{-1}(B) \in \mathcal{M} \quad \forall B \in \mathcal{B}(\mathbb{R}).$$

Definition 1.2

Let X be a metric space. A function $f : X \to \mathbb{R}$ is said to be **Borel measurable** or a **Borel function** if

$$\{x \in X : f(x) < \alpha\} \in \mathcal{B}(X) \quad \forall \alpha \in \mathbb{R}.$$

Clearly,

Every continuous function is Borel measurable.

- Likewise, if $X = \mathbb{R}^n$ and $\mathcal{M}$ consists of the Lebesgue measurable sets, $\mathcal{M} = \mathcal{L}^n$, the measurable functions are often called **Lebesgue measurable**.

Theorem 1.2

Let A be a Lebesgue measurable subset of $\mathbb{R}^n$ and $f : A \to \mathbb{R}$ be continuous. Then f is measurable with respect to n-dimensional Lebesgue measure.

In words, *a real-valued function that is continuous on its measurable domain is measurable.*

- *Not* every measurable function is continuous.

Theorem 1.3

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

Thus if f is measurable with domain A, then $|f|, |f|^p$ and f^n are measurable on A for each $p > 0$ and $n \in \mathbb{N}$.

Notation

In this chapter, we use the following short-hand notation: For numerical functions f and g on $A \in \mathcal{M}$,

$$\{f \leq g\} := \left\{x \in A : f(x) \leq g(x)\right\}.$$

The sets $\{f < g\}$, $\{f = g\}$, $\{f \neq g\}$, etc., are defined analogously. For instance, if $\alpha \in \overline{\mathbb{R}}$ then

$$\{f = \alpha\} = \left\{x \in A : f(x) = \alpha\right\},$$
$$\{f > \alpha\} = \left\{x \in A : f(x) > \alpha\right\}.$$

Theorem 1.4

For any measurable functions $f, g : A \to \overline{\mathbb{R}}$ the sets

$$\{f < g\}, \quad \{f \leq g\}, \quad \{f = g\}, \quad \{f \neq g\}$$

are measurable.

Theorem 1.5

Let f and g be measurable functions, and let c be a real number. Then cf and $f \cdot g$ are measurable, and $f+g$, $f-g$ are measurable provided $f(x)+g(x)$, $f(x)-g(x)$ is everywhere defined.

Note The multiplication is everywhere defined in $\overline{\mathbb{R}}$ and $0 \cdot (\pm\infty) = (\pm\infty) \cdot 0 = 0$.

Example 1.4 If $A_1, \ldots, A_k$ and measurable sets and $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, then the function $\sum_{i=1}^{k} \alpha_i \chi_{A_i}$ is measurable.

Example 1.5 Piecewise continuous functions on $\mathbb{R}$ are Borel-measurable.

For instance, the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2 + e^x, & \text{if } x \leq 0 \\ \sin 2x, & \text{if } x > 0 \end{cases}$$

is Borel-measurable since

$$f(x) = \left(x^2 + e^x\right)\chi_{(-\infty,0]} + \sin 2x \cdot \chi_{(0,\infty)}.$$

In general, *if f is a continuous function, or a piecewise continuous function, then f is Borel-measurable.*

Theorem 1.6

For every finitely many measurable functions $f_1, \dots, f_n$, $\min\{f_1, \dots, f_n\}$ *and* $\max\{f_1, \dots, f_n\}$ *are measurable.*

Note

$$\min\{f_1, \dots, f_n\}(x) = \min\{f_1(x), \dots, f_n(x)\},$$
$$\max\{f_1, \dots, f_n\}(x) = \max\{f_1(x), \dots, f_n(x)\},$$
$$(\lim f_n)(x) = \lim f_n(x),$$
$$(\sup f_n)(x) = \sup f_n(x),$$
$$(\inf f_n)(x) = \inf f_n(x),$$

.

Example 1.6 To every numerical function $f : A \to \overline{\mathbb{R}}$, two other functions on A are associated:

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

f^+ is called the **positive part** and f^- is called the **negative part** of f. We have

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

- f is measurable $\iff$ f^+ and f^- are measurable.
- If f is measurable, then so is $|f|$.

When considering sequences of functions $\{f_n\}$ and their convergence to a function f, we often implicitly assume that all of the functions have a common domain.

Theorem 1.7

If $\{f_n\}$ is a sequence of measurable functions, then the functions

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

are all measurable. Thus if $f(x) = \lim f_n(x)$ exists in $\overline{\mathbb{R}}$ at every x, then f is measurable.

Corollary 1.8

(a) If $f_n : A \to \mathbb{R}$ is a sequence of measurable functions and if the series $\sum_{n=1}^{\infty} f_n(x)$ exists in $\overline{\mathbb{R}}$ for all $x \in A$, then $f(x) = \sum_{n=1}^{\infty} f_n(x)$ is measurable on A.

(b) If $f_n : A \to [0, \infty]$ is a sequence of nonnegative extended real-valued measurable functions, then $f(x) = \sum_{n=1}^{\infty} f_n(x)$ is measurable on A.

Example 1.7 A finite function $f : X \to \mathbb{R}$ taking on countable values (that is, $f(X)$ is countable) is measurable if and only if

$$\{f = a\} \in \mathcal{M} \quad \text{for all } a \in \mathbb{R}.$$

Definition 1.3

For a measure space $(X, \mathcal{M}, \mu)$ and a measurable subset A of X, we say that a property holds **almost everywhere** on A (abbreviated **a.e.**), or it holds for **almost all** x in A, provided it holds on $A \setminus B$, where B is a null set.

If more precision is needed, we shall speak of a **μ-null** set or **μ-almost everywhere**, and write **μ-a.e.** or **a.e. $[\mu]$** for "almost everywhere with respect to μ."

Example 1.8 Let $(X, \mathcal{M}, \mu)$ be a measure space and let f, f_n, g be measurable functions on X. Then

$$\begin{aligned} f = g \text{ a.e.} &\iff \mu\big(\{x \in X : f(x) \neq g(x)\}\big) = 0, \\ f \leq g \text{ a.e.} &\iff \mu\big(\{x \in X : f(x) > g(x)\}\big) = 0, \\ f_n \to f \text{ a.e.} &\iff \mu\big(\{x \in X : f_n(x) \not\to f(x)\}\big) = 0. \end{aligned}$$

• In probability theory, convergence almost everywhere is known as **convergence almost surely** or **convergence with probability 1**.

Remark 1.2

- If $P(x)$ holds a.e., then the set $\{x : P(x) \text{ false}\}$ may in general *not* be measurable.
- However, if the measure μ is complete, then

$$P(x) \text{ holds a.e.} \iff \mu\big(\{x : P(x) \text{ false}\}\big) = 0.$$

Theorem 1.9

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

(a) *If f is a measurable function and if $f = g$ a.e., then g is measurable.*

(b) *If f_n is measurable for all n and $f_n \to f$ a.e., then f is measurable.*

Definition 2.1

Let f and $\{f_k\}$ be measurable functions which are defined and finite a.e. in a set A. Then $\{f_k\}$ is said to **converge in measure** on A to f if for every $\eta > 0$,

$$\lim_{n\to\infty} \mu\big(\{x \in A : |f_n(x) - f(x)| \geq \eta\}\big) = 0.$$

Note

In the theory of probability convergence in measure is referred to as **convergence in probability**.

Example 2.1

(a) The sequences $f_n = \frac{1}{n}\chi_{(0,n)}$ and $g_n = n\chi_{[0,n^{-1}]}$ converge to zero in measure.

(b) If f_n, f are real-valued measurable functions on A and $f_n \to f$ uniformly on A, then $f_n \to f$ in measure on A.

(c) If f is measurable and $\mu(A_n) \to 0$, then $f\chi_{A_n} \to 0$ in measure.

Remark 2.1

- If a sequence $\{f_n\}$ converges in measure to functions f and g, then $f = g$ almost everywhere.
- Hence up to a redefinition of functions on measure zero sets, the limit in the sense of convergence in measure is unique.

Convergence in measure vs. convergence a.e.

Theorem 2.1

Let f and f_k, $k = 1, 2, \ldots$, be measurable and finite a.e. on A. If f_k converges almost everywhere on A to f and $\mu(A) < \infty$, then f_k converges to f in measure on A.

$$\text{convergence a.e.} \overset{\mu(E)<\infty}{====>} \text{convergence in measure}$$

Example 2.2 Suppose $\{A_n\}$ is an increasing sequence with $\bigcup_{n=1}^{\infty} A_n = A$ and $\mu(A) < \infty$. If f is measurable and finite a.e. on A, then $f\chi_{A_n} \to f\chi_A$ in measure.

Theorem 2.2

If $\{f_n\}$ converges to f in measure on A, then there is a subsequence $\{f_{n_k}\}$ that converges pointwise a.e. on A to f.

$$\text{convergence in measure} \xRightarrow{\textit{subsequence}} \text{convergence a.e.}$$

From now on $(X, \mathcal{M}, \mu)$ is assumed to be a fixed measure space and $A \in \mathcal{M}$.

Definition 3.1

A real-valued function on A is called a **simple** function if it is measurable and assumes only finitely many different values.

Note We do *not* allow simple functions to assume the values $\pm\infty$.

Remark 3.1 If φ and ψ are simple functions on X and $c \in \mathbb{R}$, then

$$c\varphi, \quad \varphi{\cdot}\psi, \quad \varphi{\pm}\psi, \quad \max\{\varphi, \psi\}, \quad \text{and} \quad \min\{\varphi, \psi\}$$

are also simple functions.

In worlds, finite sums, finite products, and finite suprema and infima of simple functions are again simple functions.

In particular, if φ is simple then so are φ^+ and φ^-.

If s is a simple function on X assuming the distinct values $a_1, \ldots, a_n$, then the sets

$$A_i = \{x \in X : s(x) = a_i\}$$

are all measurable and pairwise disjoint, and

$$s(x) = \sum_{i=1}^{n} a_i \chi_{A_i}(x) \quad \forall x \in X.$$

This expression is called the **standard representation** of s.

Theorem 3.1

If $f : A \to \overline{\mathbb{R}}$ is nonnegative and measurable, then there is a sequence $\{s_n\}$ of simple functions on A such that

(a) $0 \leq s_1 \leq s_2 \leq \cdots \leq f$, $s_n \to f$ *pointwise*, *and*

(b) $s_n \to f$ *uniformly on any set on which f is bounded.*

In this section we will define $\int_X f d\mu$ for a class of measurable functions.

This is a three step procedure:

1. Integration of simple functions,
2. Integration of nonnegative functions,
3. Integration of general functions.

This sequence of three steps is also useful in proving integration formulas.

The integral of non-negative simple functions

- $\int_a^b 1dx = (b-a)$.
- If $[c,d] \subset [a,b]$, then $\int_a^b \chi_{[c,d]}dx = (d-c)$.
- $\int_X \chi_A d\mu = \mu(A)$.
- $\int_X \left(\sum_{k=1}^m c_k\chi_{A_k}\right) dx = \sum_{k=1}^m c_k\mu(A_k)$.

Definition 3.2

Let s be a nonnegative simple function with standard representation $s = \sum_{i=1}^{n} a_i \chi_{A_i}$, that is, $a_1, \ldots, a_n$ are the distinct values of s, and

$$A_i = \{x \in X : s(x) = a_i\}.$$

Then the **μ-integral** of s (or simply, the **integral** of s) is defined by

$$\int_X s d\mu = \sum_{i=1}^{n} a_i \mu(A_i).$$

The convention $0 \cdot \infty = 0$ is used here.

Lemma 3.2

Let $s = \sum_{i=1}^{m} a_i \chi_{A_i}$ be the standard representation of a nonnegative simple function s and let $s = \sum_{j=1}^{n} b_j \chi_{B_j}$ be another representation in which the the B_j are disjoint and measurable. Then

$$\int_X s d\mu = \sum_{j=1}^{n} b_j \mu(B_j).$$

In other words, the integral of a nonnegative simple function does not depend upon its particular representation.

Definition 3.3

If $A \in \mathcal{M}$ we define

$$\int_A s d\mu := \int_X s \cdot \chi_A d\mu.$$

- Note that for any nonnegative simple function s,

$$0 \leq \int_A s d\mu \leq \infty,$$

that is, $\int_A s d\mu$ is a nonnegative extended number.

- If
$$s = \sum_{i=1}^{n} a_i \chi_{A_i},$$
where $A_i \in \mathcal{M}$ and $A_i \cap A_j = \emptyset$ for $i \neq j$, then
$$\int_A s d\mu = \sum_{i=1}^{n} a_i \mu(A_i \cap A).$$

Particular cases: If $A, E \in \mathcal{M}$, then

$$\boxed{\int_X \chi_E d\mu = \mu(E) \quad \text{and} \quad \int_A \chi_E d\mu = \mu(A \cap E)}$$

It is sometimes convenient to display the argument of s explicitly, especially when $s(x)$ is given by a formula in terms of x or when there are other variables involved; in this case we shall use the notation

$$\int_A s(x) d\mu(x).$$

Lemma 3.3

Let φ and ψ be nonnegative simple functions on X.

(a) *If $c \geq 0$, then $\int_A c\varphi d\mu = c\int_A \varphi d\mu$.*

(b) $\int_A (\varphi + \psi) d\mu = \int_A \varphi d\mu + \int_A \psi d\mu$.

(c) *If $\varphi \leq \psi$ on A, then $\int_A \varphi d\mu \leq \int_A \psi d\mu$.*

(d) *The function $A \mapsto \int_A \varphi d\mu$ is a measure on $\mathcal{M}$.*

Definition 3.4 (The integral of nonnegative measurable functions)

If f is a nonnegative measurable function on X, we define the **μ-integral** of f on the set X with respect to the measure μ to be

$$\begin{aligned}\int_X f\,d\mu &= \int_X f(x)d\mu(x)\\ &= \sup\left\{\int_X s\,d\mu : 0 \le s \le f,\ s \text{ simple}\right\}.\end{aligned}$$

$\int_X f\,d\mu$ is also called the **Lebesgue integral** of f over X, with respect to the measure μ.

Remark 3.2

- If f is a nonnegative measurable function on X, $\int_X f d\mu$ always exists. It is a number in $[0, \infty]$ and may equal ∞.
- When f is simple, the two definitions of $\int_X f d\mu$ agree.

For a general measurable function f, the functions f^+ and f^- are nonnegative and are measurable, and consequently,

$$\int_X f^+ d\mu \quad \text{and} \quad \int_X f^- d\mu$$

are well-defined extended real numbers in $[0, \infty]$.

Definition 4.1

If f is measurable on X, we define

$$\int_X f\,d\mu = \int_X f^+ d\mu - \int_X f^- d\mu,$$

provided that the two integrals on the right are not both ∞.

The **integral of f over a measurable set A** with respect to μ is defined to be

$$\int_X (f \cdot \chi_A)\,d\mu$$

and denoted by $\int_A f\,d\mu$.

We say that f is **μ-integrable** on A (or simply **integrable** if μ is understood) if $\int_A f\,d\mu$ is finite.

Note

- If $\mu = m_n$, traditionally one writes

$$\int_A f dx \quad \text{or} \quad \int_A f(x)dx \qquad \text{for} \qquad \int_A f dm_n.$$

- Since $\chi_\emptyset = 0$, we have

$$\int_\emptyset f\, d\mu = 0$$

for all measurable functions f.

- If $(X, \mathcal{M}, \mu)$ is a probability space and $\xi : X \to \mathbb{R}$ is a random variable (that is, ξ is measurable), then $\int_X \xi d\mu$ is called the **expectation** of ξ and is denoted by $E(\xi)$.

Remark 3.3 • If f is a measurable function and A is a measurable set, then

$$\int_A f d\mu = \int_A f^+ d\mu - \int_A f^- d\mu$$

provided that at least one of $\int_A f^+ d\mu$ and $\int_A f^- d\mu$ is finite.

• Clearly,

$$\begin{aligned} f \text{ integrable on } A &\iff \int_A f\,d\mu \text{ finite} \\ &\iff \int_A f^+ d\mu \text{ and } \int_A f^- d\mu \text{ finite} \end{aligned}$$

Example 3.1 (**Chebychev's Inequality**)
If f is a nonnegative measurable function on A and c a positive real number, then

$$\mu(\{f \geq c\}) \leq \frac{1}{c} \int_A f d\mu.$$

Example 3.2 If $f(x) \equiv c$ is a constant function with $c \in \overline{\mathbb{R}}$ and A is measurable, then

$$\int_A c d\mu = c\mu(A).$$

Thus if $c \neq 0$ is a finite value, then the constant function $f(x) \equiv c$ is integrable over A if and only if $\mu(A) < \infty$.

Passage of the limit under the integral sign

We now establish a criterion for justifying **passage of the limit under the integral sign**, that is,

$$\lim_{n\to\infty}\int_A f_n d\mu = \int_A \left(\lim_{n\to\infty} f_n\right) d\mu.$$

Theorem 3.4 (The Monotone Convergence Theorem)

If $\{f_n\}$ is an increasing sequence of nonnegative measurable functions on A and $f = \lim_{n\to\infty} f_n$ on A, then

$$\lim_{n\to\infty} \int_A f_n d\mu = \int_A f d\mu.$$

If $\{f_n\}$ is an increasing sequence of functions and $f = \lim_{n\to\infty} f_n$, then we denote $f_n \nearrow f$.

Theorem 3.4 can be written as

$$0 \le f_n \nearrow f \text{ on } A \Longrightarrow \int_A f_n d\mu \nearrow \int_A f d\mu.$$

Note The following result is derived from Theorems 3.4 and 3.1

Let f be a nonnegative measurable function on A. Then there is an increasing sequence $\{\varphi_n\}$ of simple functions that converges pointwise on A to f and

$$\lim_{n\to\infty}\int_A \varphi_n d\mu = \int_A f d\mu.$$

We now establish basic properties of the Lebesgue integral, including properties that indicate how Lebesgue integration interacts with passages to the limit.

Theorem 4.1 (Monotonicity)

Let f and g be measurable functions. If their integrals over A are defined, and if $f \leq g$ on A, then $\int_A f\,d\mu \leq \int_A g d\mu$.

Theorem 4.2

(a) *If f is measurable and $f = 0$ a.e. on A, then $\int_A f d\mu = 0$.*

(b) *If f is a measurable function and $\mu(A) = 0$, then $\int_A f d\mu = 0$.*

(c) *If A and B are measurable sets with $B \subset A$ and if f is a nonnegative measurable function, then $\int_B f d\mu \le \int_A f d\mu$.*

(d) *If A and B are measurable sets with $B \subset A$ and $\int_A f d\mu$ is defined, then so is $\int_B f d\mu$.*

Note

- If A and B are measurable sets with $B \subset A$ and $\int_A f d\mu$ is defined, then

$$\int_B f d\mu = \int_A f \chi_B d\mu.$$

- If f is integrable over A, then f is integrable over every measurable subset of A.

Theorem 4.3 (Linearity)

(a) *If $\int_A f\,d\mu$ is defined and if c is a real constant, then $\int_A (cf)\,d\mu$ is defined and*

$$\int_A (cf)\,d\mu = c\int_A f\,d\mu.$$

(b) *Let f and g are measurable functions on A and assume that $f+g$ is everywhere defined. If $\int_A f\,d\mu$, $\int_A g\,d\mu$ exist, and $\int_A f\,d\mu + \int_A g\,d\mu$ is defined, then*

$$\int_A (f+g)\,d\mu = \int_A f\,d\mu + \int_A g\,d\mu.$$

Remark 4.1

(a) If both f and g are nonnegative and measurable on A, then

$$\int_A (f+g)d\mu = \int_A f d\mu + \int_A g d\mu.$$

(b) If f is measurable on A, then

$$\int_A |f| d\mu = \int_A f^+ d\mu + \int_A f^- d\mu.$$

Corollary 4.4

If $\int_A f\,d\mu$ is defined, then $\left|\int_A f\,d\mu\right| \leq \int_A |f|\,d\mu$.

Corollary 4.5

(a) If f is measurable, nonnegative on A, and $\int_A f\,d\mu = 0$, then $f = 0$ a.e. on A.

(b) If f is measurable, nonpositive on A, and $\int_A f\,d\mu = 0$, then $f = 0$ a.e. on A.

Corollary 4.6

(a) *Let f, g be measurable functions on A. If $f = g$ a.e. on A and $\int_A f d\mu$ exists, then so does $\int_A g d\mu$ and*

$$\int_A f d\mu = \int_A g d\mu.$$

(b) *If f is a measurable function, A a measurable set, and $\mu(B) = 0$, then*

$$\int_{A\cup B} f d\mu = \int_{A\setminus B} f d\mu = \int_A f d\mu$$

provided that one of the integrals exists.

Corollary 4.7

For every sequence $\{f_n\}$ of nonnegative measurable functions, $\sum_{n=1}^{\infty} f_n$ is measurable and

$$\int_A \left(\sum_{n=1}^{\infty} f_n \right) d\mu = \sum_{n=1}^{\infty} \int_A f_n d\mu.$$

This corollary shows that for the purposes of integration it makes no difference if we alter functions on null sets (that is, $\int_A f d\mu$ is unchanged if we modify f in a set of measure zero).

Thus in any integration theorem, we may freely use the phrase "almost everywhere."

Theorem 4.8

Let f be measurable on A. f is integrable on A if and only if $|f|$ is integrable on A.

$$f \text{ is integrable on } A \Longleftrightarrow \begin{cases} f \text{ is measurable on } A \text{ and} \\ |f| \text{ is integrable on } A \end{cases}$$

Corollary 4.9

Let f be a measurable function on A.

(a) *If g is integrable over A and dominates f on A in the sense that $|f| \leq g$ a.e. on A, then f is integrable over A.*

(b) *In particular, if $\mu(A)$ is finite and there is a real number c such that $|f| \leq c$ a.e., then f is integrable on A.*

Theorem 4.10

If f is integrable on A, then f is finite a.e. on A.

Example 4.1 (**The first Borel-Cantelli lemma**)

Let $\{A_k\}$ be a sequence of measurable sets in X, such that

$$\sum_{k=1}^{\infty} \mu(A_k) < \infty.$$

Then almost all $x \in X$ lie in at most finitely many of the sets A_k.

Example 4.2 If f and g are integrable functions on A, then the functions

$$\max\{f, g\} \quad \text{and} \quad \min\{f, g\}$$

are integrable on A.

Remark 4.2 If $\int_A f d\mu$ exists and g is integrable over A, then $f \pm g$ are defined a.e. on A and we have

$$\int_A (f \pm g) d\mu = \int_A f d\mu \pm \int_A g d\mu.$$

Theorem 4.11 (σ-additivity)

Let f be a fixed measurable function, and suppose that $\int_X f\,d\mu$ is defined. Then the set function

$$\nu(A) = \int_A f\,d\mu, \quad A \in \mathcal{M}$$

is σ-additive. In paricular, if $f \geq 0$ a.e, then ν is a measure on $\mathcal{M}$.

Example 4.3 Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. Recall that a **random variable** ξ is a measurable function from the space Ω to $\mathbb{R}$. The **expectation** (or **mean value**) of ξ is defined by

$$E(\xi) = \int_\Omega \xi d\mathbf{P}$$

provided the integral exists.

We say that ξ is **discrete** if the set of values of ξ is countable.

Show that if ξ is a discrete random variable and $\xi(\Omega) = \{x_1, x_2, \ldots\}$ ($x_i \neq x_j$ for $i \neq j$), then

$$E(\xi) = \sum_k x_k \mathbf{P}(\{\xi = x_k\})$$

whenever this sum converges absolutely, in that

$$\sum_k |x_k| \mathbf{P}(\{\xi = x_k\}) < \infty.$$

Corollary 4.12 (Additivity)

Suppose that $\int_X f\,d\mu$ *is defined. If* $A_1, \dots, A_n$ *are disjoint measurable subsets of* X*, then*

$$\int_{\bigcup_{k=1}^n A_k} f\,d\mu = \sum_{k=1}^n \int_{A_k} f\,d\mu.$$

Theorem 4.13

If f is an integrable function on X, then for any $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\int_A |f| d\mu < \epsilon \quad \text{whenever} \quad \mu(A) < \delta.$$

$$\left(\int_X |f| d\mu < +\infty \right)$$

$$\Downarrow$$

$$(\forall \epsilon > 0)(\exists \delta > 0)\big(\forall \mu(A) < \delta\big)\left(\int_A |f| d\mu < \epsilon \right).$$

Lemma 4.14 (Fatou)

If A is a measurable set and if $\{f_n\}$ is a sequence of nonnegative measurable functions, then

$$\int_A \left(\liminf_{n\to\infty} f_n\right) d\mu \leq \liminf_{n\to\infty} \int_A f_n d\mu.$$

In particular, if $f_n \to f$ a.e., then

$$\int_A f\, d\mu \leq \liminf_{n\to\infty} \int_A f_n d\mu.$$

Example 4.4 Show that if f_n are nonnegative measurable functions on A, $f_n \to f$ and $f \geq f_n$ a.e for all n, then

$$\lim_{n\to\infty} \int_A f_n d\mu = \int_A f d\mu.$$

Theorem 4.15 (Dominated Convergence Theorem)

Let f, f_n be measurable functions such that

(a) $f_n \to f$ *a.e. on* A;

(b) *there exists an integrable function* g *on* A *such that* $|f_n| \leq g$ *a.e. for all* n.

Then the function f *is integrable on* A *and*

$$\lim_{n\to\infty} \int_A f_n d\mu = \int_A f d\mu.$$

Corollary 4.16

Let A be a set of finite measure, let c be a positive number, and suppose that $\{f_n\}$ is a sequence of measurable functions such that $|f_n| \leq c$ a.e. for all n. If $f_n \to f$ a.e., then $\lim \int_A f_n d\mu$ exists, f is integrable on A, and

$$\int_A f\, d\mu = \lim_{n\to\infty} \int_A f_n d\mu.$$

Corollary 4.17

Suppose that $\{f_n\}$ is a sequence of integrable functions on A such that

$$\sum_{n=1}^{\infty} \int_A |f_n| d\mu < \infty.$$

Then the series $\sum_{n=1}^{\infty} f_n$ converges a.e. to an integrable function and

$$\int_A \left(\sum_{n=1}^{\infty} f_n \right) d\mu = \sum_{n=1}^{\infty} \int_A f_n d\mu.$$

3.5 THE RIEMANN INTEGRAL AS A LEBESGUE INTEGRAL

At this point it is appropriate to study the relation between the Lebesgue and Riemann integrals on $\mathbb{R}^n$.

In this section, "f is Lebesgue integrable" means that f is integrable with respect to the Lebesgue measure.

Theorem 5.1

If a function f is Riemann integrable on the closed, bounded interval $[a, b]$, then it is Lebesgue integrable on $[a, b]$ and the two integrals are equal.

$$\int_{[a,b]} f\, dm = \int_a^b f(x)dx.$$

In view of Theorem 5.1 we sometimes use the notation

$$\int_a^b f(x)dx$$

for Lebesgue integral on $\mathbb{R}$.

Theorem 5.2

A bounded function f defined on the closed, bounded interval $[a, b]$ is Riemann integrable if and only if it is continuous almost everywhere.

For example, the Dirichlet function $D(x)$ defined by

$$D(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

is *not* Riemann integrable on $[0, 1]$ but

$$\int_0^1 D(x)dm = 0.$$

Theorem 5.3

Suppose that f and $|f|$ are integrable on an interval I (bounded or unbounded) in the improper Riemann sense. Then f is Lebesgue integrable on I and its improper Riemann integral equals its Lebesgue integral.

Remark 5.1 Theorems 5.1–5.3 are valid for Riemann integral of multivariable functions.

Example 5.1 Recall that

$$\sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}.$$

Evaluate the integral

$$\int_0^1 \left(\frac{\ln x}{1-x}\right)^2 dx.$$

ANS. $\frac{\pi^2}{3}$; Note that $\int_0^1 nx^{n-1}(\ln x)^2 dx = \frac{2}{n^2}$.

Theorem 5.4

Let $(X, \mathcal{M}, \mu)$ be a measure space and let $f : X \times (a, b) \to \mathbb{R}$ be a function such that for every $t \in (a, b)$ the function $x \mapsto f(x, t)$ is integrable.

(a) *Suppose that for almost x the function $t \mapsto f(x, t)$ is continuous and there exists a μ-integrable function g such that for each fixed $t \in (a, b)$ we have $|f(x, t)| \leq g(x)$ almost everywhere. Then the function*
$$J(t) = \int_X f(x, t)d\mu(x)$$
is continuous.

Theorem 5.4 (cont'd)

(b) *Suppose that, for almost x the function $t \mapsto f(x,t)$ is differentiable and there exists a μ-integrable function $g(x)$ such that for almost x we have $|\partial f(x,t)/\partial t| \leq g(x)$ for all t simultaneously. Then, the function J is differentiable and*

$$J'(t) = \int_X \frac{\partial f}{\partial t}(x,t)d\mu(x), \quad t \in (a,b).$$

3.6 PRODUCT MEASURES AND ITERATED INTEGRALS

Throughout this section $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two reference measure spaces.

Consider the Cartesian product $X \times Y$ of X and Y. If $A \subset X$ and $B \subset Y$, we call $A \times B$ a **rectangle**. If $A \in \mathcal{A}$ and $B \in \mathcal{B}$, we call $A \times B$ a **measurable rectangle**. We have

$$\begin{aligned}(A \times B) \cap (E \times F) &= (A \cap E) \times (B \cap F),\\ (A \times B)^c &= (X \times B^c) \cup (A^c \times B).\end{aligned}$$

Suppose $A \times B$ is a rectangle that is a countable disjoint union of measurable rectangles $A_i \times B_i$. Then

$$\mu(A) \cdot \nu(B) = \sum_i \mu(A_i) \cdot \nu(B_i).$$

The collection $\mathcal{C}$ of finite disjoint unions of measurable rectangles is an algebra over $X \times Y$.

If $E \in \mathcal{C}$ is the disjoint union of measurable rectangles $A_1 \times B_1, \ldots, A_n \times B_n$, we set

$$\lambda(E) = \sum_{i=1}^{n} \mu(A_i) \cdot \nu(B_i)$$

with the usual convention that $0 \cdot \infty = 0$. Then λ is well defined on $\mathcal{C}$ and λ is a premeasure on $\mathcal{C}$.

According to Theorem 3.3 of Chapter 2, λ can be extended to an outer measure on $X \times Y$.

Definition 6.1

Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces, $\mathcal{C}$ the algebra of disjoint unions of measurable rectangles contained in $X \times Y$, and λ the premeasure defined on $\mathcal{C}$ by

$$\lambda(E) = \sum_{i=1}^{n} \mu(A_i) \cdot \nu(B_i)$$

if $E \in \mathcal{C}$ is the disjoint union of measurable rectangles $A_1 \times B_1, \ldots, A_n \times B_n$. By the **product measure** $\mu \times \nu$ of μ and ν we mean the Carathéodory extension of $\lambda : \mathcal{C} \to [0, \infty]$ defined on the σ-algebra of λ^*-measurable subsets of $X \times Y$.

Let E be a subset of $X \times Y$ and f a function on E. For a point $x \in X$, we call the set

$$E_x = \{y \in Y : (x, y) \in E\} \subset Y$$

the ***x*-section** of E and the function $f(x, \cdot)$ defined on E_x by $f(x, \cdot)(y) = f(x, y)$ the ***x*-section** of f.

Similarly, for a point $y \in Y$, the set

$$E^y = \{x \in X : (x, y) \in E\} \subset X$$

is called the ***y*-section** of E and the function $f(\cdot, y)$ defined on E^y by $f(\cdot, y)(x) = f(x, y)$ is called the ***y*-section** of f.

Our goal now is to determine what is necessary in order that

$$\int_{X\times Y} f\, d(\mu\times\nu) = \int_X \left[\int_Y f(x,y)d\nu(y)\right] d\mu(x) = \int_Y \left[\int_X f(x,y)d\mu(x)\right] d\nu(y). \tag{1}$$

This is called **iterated integration**.

Theorem 6.1 (Tonelli)

Suppose that $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are σ-finite measure spaces. Let f be a nonnegative $(\mu \times \nu)$-measurable function on $X \times Y$ then the functions

$$g(x) = \int_Y f(x, \cdot) d\nu \quad \text{and} \quad h(y) = \int_X f(\cdot, y) d\mu$$

are measurable, and (1) holds.

Theorem 6.2 (Fubini)

Suppose that $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are σ-finite measure spaces. If f is integrable over $X \times Y$ with respect to the product measure $\mu \times \nu$, then

(i) *the function $f(x, \cdot)$ is integrable over Y with respect to ν for a.e. $x \in X$, $f(\cdot, y)$ is integrable over X with respect to μ for a.e. $y \in Y$;*

(ii) *the functions $g(x) = \int_Y f(x, \cdot)d\nu$ and $h(y) = \int_X f(\cdot, y)d\mu$ are integrable, and (1) holds.*

The Fubini and Tonelli theorems are usually referred to as "the method of computing a double integral by changing the order of integration."

Remark 6.1 We shall usually omit the brackets in the iterated integrals in (1), thus,

$$\int \left[\int f(x,y) d\mu(x) \right] d\nu(y) = \iint f(x,y) d\mu(x) d\nu(y) = \iint f \, d\mu d\nu.$$

The principal theme of the remain sections is the concept of differentiating a measure ν with respect to another measure μ on the same σ-algebra.

To do this, it is useful to generalize the notion of measure so as to allow measures to assume negative values.

From now on, $(X, \mathcal{M})$ is a measurable space. All sets involved are assumed as usual to lie in $\mathcal{M}$.

If μ_1 and μ_2 are two measures defined on the same measurable space $(X, \mathcal{M})$, then

$$\mu(A) = \mu_1(A) + \mu_2(A), \quad A \in \mathcal{M}$$

is a measure on $\mathcal{M}$.

Question:
How about the set function

$$\nu(A) = \mu_1(A) - \mu_2(A)$$

if it is defined for every $A \in \mathcal{M}$?

Definition 7.1

A set function $\mu : \mathcal{M} \to \overline{\mathbb{R}}$ is said to be a **signed measure** if it satisfies the following properties:

(a) μ assumes at most one of the values $+\infty$, $-\infty$,

(b) $\mu(\emptyset) = 0$, and

(c) μ is σ-additive, that is, if $\{A_n\}$ is a disjoint sequence of members of $\mathcal{M}$, then

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

Note In (c) if $\mu\left(\cup_{n=1}^{\infty} A_n\right)$ is finite, then the series $\sum_{n=1}^{\infty} \mu(A_n)$ converges absolutely.

In practice, one usually deals with finite signed measures.

Clearly every measure is a signed measure; for emphasis we shall sometimes refer to measures as **positive measures**.

Example 7.1 Let f be a function such that $\int_X f d\mu$ is defined. Then the set function

$$\nu(A) := \int_A f d\mu, \quad A \in \mathcal{M},$$

is a signed measure. We call ν the **indefinite integral** of f (with respect to μ).

Example 7.2 A measure is a special case of a signed measure. Conversely, if μ and ν are measures on a σ-algebra $\mathcal{M}$, at least one of which is finite, then $\varphi(A) = \mu(A) - \nu(A)$ is a signed measure on $\mathcal{M}$.

Theorem 7.1

Let μ be a signed measure on $\mathcal{M}$.

(a) If $\{A_n\}$ is an increasing sequence in $\mathcal{M}$, then $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$.

(b) If $\{A_n\}$ is a decreasing sequence in $\mathcal{M}$ and $\mu(A_1)$ is finite, then

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

Note that signed measures *need not be monotone* unless they are positive measures.

Definition 7.2

A measurable set A is said to be a **positive set** with respect to a signed measure ν if $\nu(E) \geq 0$ for all $E \in \mathcal{M}$ with $E \subset A$. Similarly, a set B is called **negative** (respectively, **null**) for ν provided it is measurable and every measurable subset of B has nonpositive (respectively, zero) ν measure.

$$A \text{ is positive} \iff (\forall E \subset A,\ E \in \mathcal{M})\big(\nu(E) \geq 0\big)$$
$$B \text{ is negative} \iff (\forall E \subset B,\ E \in \mathcal{M})\big(\nu(E) \leq 0\big)$$
$$C \text{ is null} \iff (\forall E \subset C,\ E \in \mathcal{M})\big(\nu(E) = 0\big).$$

Example 7.3 Suppose that $\int_X f d\mu$ exists and

$$\nu(E) := \int_E f d\mu, \quad E \in \mathcal{M}.$$

A set $A \in \mathcal{M}$ is positive, negative, or null for ν precisely when $f \geq 0$, $f \leq 0$, or $f = 0$ μ-a.e. on A.

Remark 7.1

- If the sets A_n are positive, then $A = \bigcup_n A_n$ is also positive.
- The conclusions remain valid if positive is replaced by negative.

Theorem 7.2 (Hahn Decomposition)

Let ν be a signed measure on the measurable space $(X, \mathcal{M})$. Then there is a positive set P for ν and a negative set N for ν for which

$$X = P \cup N \quad \text{and} \quad P \cap N = \emptyset.$$

A decomposition of X into the union of two disjoint sets, $X = P \cup N$, for which P is positive for ν and N negative is called a **Hahn decomposition** for ν.

ν is a signed measure on $(X, \mathcal{M})$

$$\Downarrow$$

$$\exists P, N \in \mathcal{M}: \; X = P \cup N, \; P \cap N = \emptyset,$$

P is a positive set and N is a negative set for ν

Example 7.4 If $\int_X f d\mu$ is defined and

$$\nu(A) = \int_A f d\mu,$$

then we can take

$$P = \{f \geq 0\} \quad \text{and} \quad N = \{f < 0\}.$$

Note Hahn decomposition for ν is *not* unique.

Definition 7.3

Two measures ν_1 and ν_2 on $(X, \mathcal{M})$ are said to be **mutually singular** (in symbols $\nu_1 \perp \nu_2$) if there are disjoint measurable sets A and B with $X = A \cup B$ for which $\nu_1(A) = \nu_2(B) = 0$. In this case ν_1 is also said to be **singular with respect to** ν_2 and ν_2 singular with respect to ν_1

$$\nu_1 \perp \nu_2 \iff \exists A : \; \nu_1(A) = \nu_2(A^c) = 0.$$

Example 7.5

(a) Let $(X, \mathcal{M}) = (\mathbb{R}, \mathcal{L})$. For any $a \in \mathbb{R}$, $m \perp \delta_a$.

(b) Set, for $A \in \mathcal{L}$,

$$\mu(A) = m\big(A \cap (-\infty, 0]\big)$$

and

$$\nu(A) = \int_{A\cap(0,\infty)} e^{-x} dx.$$

Then $\mu\big((0,\infty)\big) = 0 = \nu\big((-\infty, 0]\big)$. Thus $\mu \perp \nu$.

Theorem 7.3 (The Jordan Decomposition Theorem)

Let ν be a signed measure on the measurable space $(X, \mathcal{M})$. Then there are two mutually singular measures ν^+ and ν^- on $(X, \mathcal{M})$ for which

$$\nu = \nu^+ - \nu^-.$$

Moreover, there is only one such pair of mutually singular measures.

$\nu = \nu^+ - \nu^-$ is called the **Jordan decomposition** of ν.

ν is a signed measure on $(X, \mathcal{M})$

$$\Downarrow$$

$\exists!$ a pair of mutually singular measures ν^+ and ν^-:

$$\nu = \nu^+ - \nu^-$$

If A, B is a Hahn decomposition for ν, then

$$\nu^+(E) = \nu(E \cap A),\ \nu^-(E) = -\nu(E \cap B), \quad E \in \mathcal{M}.$$

The measures ν^+ and ν^- are called the **positive** and **negative parts** (or **variations**) of ν.

Note that ν^+, ν^- do not depend on the particular Hahn decomposition chosen.

The measure $|\nu|$ is defined on $\mathcal{M}$ by

$$|\nu|(E) = \nu^+(E) + \nu^-(E), \quad E \in \mathcal{M}.$$

$|\nu|$ is called the **total variation** of ν.

Example 7.6 Suppose that $\int_X f d\mu$ exists, define $P = \{f \geq 0\}$, $N = \{f < 0\}$, and

$$\nu(A) = \int_A f d\mu, \quad A \in \mathcal{M}.$$

Then $\{P, N\}$ is a Hahn decomposition of X with respect to the signed measure ν,

$$\nu^+(A) = \int_A f^+ d\mu, \quad \nu^-(A) = \int_A f^- d\mu,$$

and

$$|\nu|(A) = \int_A |f| d\mu.$$

Let $(X, \mathcal{M}, \mu)$ be a measure space. Suppose that $\int_X f d\mu$ exists. Define the set function ν on $\mathcal{M}$ by

$$\nu(A) = \int_A f d\mu \quad \text{for all } A \in \mathcal{M}.$$

Then ν is a signed measure on the measurable space $(X, \mathcal{M})$ and ν has the property that

$$\mu(A) = 0 \Longrightarrow \nu(A) = 0.$$

Definition 8.1

Let $(X, \mathcal{M})$ be a measurable space. Suppose that ν is a signed measure and μ is a positive measure on $\mathcal{M}$, we say that ν is **absolutely continuous** with respect to μ, written $\nu \ll \mu$, if

$$\nu(A) = 0 \quad \text{whenever} \quad \mu(A) = 0.$$

$$\boldsymbol{\nu \ll \mu \overset{\text{def}}{\Longleftrightarrow} \left[\mu(A) = 0 \Longrightarrow \nu(A) = 0\right]}$$

Example 8.1 If μ is the counting measure and ν is an arbitrary signed measure on the measurable space $(X, \mathcal{M})$, then $\nu \ll \mu$.

Example 8.2 Let $X = \mathbb{N} \cup \{0\} = \{0, 1, 2, \ldots\}$ denote the set of all nonnegative integers. Let μ be the counting measure on $\mathcal{P}(X)$, and ν be the Poisson distribution with parameter $0 < \lambda < \infty$, i.e.,

$$\nu(A) = \sum_{k \in A} \frac{\lambda^k}{k!} e^{-\lambda}, \quad A \in \mathcal{P}(X).$$

Then $\mu \ll \nu$ and $\nu \ll \mu$.

Example 8.3 Let $(X, \mathcal{M}, \mu)$ be a measure space and let $\int_X f\, d\mu$ exists. Let ν be the signed measure defined by

$$\nu(A) = \int_A f\, d\mu, \quad A \in \mathcal{M}.$$

Then $\nu \ll \mu$.

Example 8.4 For any $a \in \mathbb{R}$ the Dirac measure δ_a on $\mathcal{L}$ is not absolutely continuous with respect to Lebesgue measure m and conversely, m is not absolutely continuous with respect to δ_a either.

Theorem 8.1 (The Lebesgue-Radon-Nikodym Theorem)

Let $(X, \mathcal{M})$ be a measurable space and let μ and ν be σ-finite measures on $(X, \mathcal{M})$. There exist unique σ-finite measures ν_a, ν_s on $(X, \mathcal{M})$ such that

$$\nu = \nu_a + \nu_s, \qquad \nu_a \ll \mu, \qquad \nu_s \perp \mu.$$

Moreover, there is a nonnegative measurable function $f : X \to \overline{\mathbb{R}}$ such that

$$\nu_a(A) = \int_A f d\mu \quad \text{for all } A \in \mathcal{M},$$

and any two such functions are equal μ-a.e.

μ, ν are σ-finite positive measures

$$\Downarrow$$

$\exists!$ a pair σ-finite positive measures ν_a, ν_s :

$$\nu = \nu_a + \nu_s, \quad \nu_a \ll \mu, \quad \nu_s \perp \mu,$$

and $\exists\, f \geq 0$:

$$\nu_a(A) = \int_A f d\mu \quad \forall\, A \in \mathcal{M}.$$

The decomposition

$$\nu = \nu_a + \nu_s, \quad \nu_a \ll \mu \quad \text{and} \quad \nu_s \perp \mu,$$

is called the **Lebesgue decomposition** of ν with respect to μ.

Corollary 8.2 (The Radon-Nikodym Theorem)

Let $(X, \mathcal{M}, \mu)$ *be a* σ*-finite measure space, and let* ν *be a* σ*-finite signed measure on* $\mathcal{M}$ *that is absolutely continuous with respect to* μ*. Then there exists a measurable function* f *such that*

$$\nu(A) = \int_A f d\mu \quad \text{for all } A \in \mathcal{M}. \tag{2}$$

Moreover, f *is unique up to a set of* μ*-measure* 0 *and if* ν *is a positive measure, then* f *is nonnegative on* X.

The function f in (2) is called the **Radon-Nikodym derivative** or **density** of ν with respect to μ and is written

$$f = \frac{d\nu}{d\mu}.$$

- For instance, let E be a measurable set. The set function ν defined by

$$\nu(A) = \mu(E \cap A), \quad A \in \mathcal{M},$$

is absolutely continuous with respect to μ and

$$\frac{d\nu}{d\mu} = \chi_E.$$

μ σ-finite measure, ν σ-finite signed on $\mathcal{M}$, $\nu \ll \mu$

$$\Downarrow$$

$\exists\, f$: $\nu(A) = \int_A f d\mu$ for all $A \in \mathcal{M}$ and

f is unique up to a set of μ-measure 0.

Example 8.5 Let $f : \mathbb{R} \to \mathbb{R}$ be an increasing and continuously differentiable function. Let ν be the Lebesgue-Stieltjes measure generated by f. By the fundamental theorem of calculus,

$$\nu\big((a, b]\big) = f(b) - f(a) = \int_a^b f'(t)dt.$$

Thus $\nu \ll m$ and the Radon-Nikodym derivative of f is precisely the ordinary derivative of f:

$$\frac{d\nu}{dm}(t) = f'(t), \quad t \in \mathbb{R}.$$

Definition 8.2

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. The **conditional expectation** of an integrable random variable ξ relative to a sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$ is any $\mathcal{G}$-measurable, integrable random variable η such that

$$\int_A \eta d\mathbf{P} = \int_A \xi d\mathbf{P} \quad \text{for every } A \in \mathcal{G}.$$

Example 8.6 Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and $\mathcal{G} = \{\emptyset, \Omega\}$.

For an integrable random variable ξ and $c := \int_\Omega \xi d\mathbf{P}$, the constant function

$$\eta(\omega) = c \quad \text{for all} \quad \omega \in \Omega$$

is a conditional expectation of ξ relative to $\mathcal{G}$.

Theorem 8.2

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and let $\mathcal{G} \subset \mathcal{F}$ be a σ-algebra. Then every random variable ξ with finite expectation has a conditional expectation relative to $\mathcal{G}$.

Note that the conditional expectation of a random variable ξ relative to $\mathcal{G}$ is the Radon-Nikodym derivative

$$\frac{d\nu}{d\mu},$$

where $\nu(A) = \int_A \xi d\mathbf{P}$, $(A \in \mathcal{G})$, and μ is the restriction of measure $\mathbf{P}$ to $\mathcal{G}$.