Chapter 4
The Integral

Based on the notions of measure spaces and measurable maps, we introduce the integral of a measurable map with respect to a general measure. This generalizes the Lebesgue integral that can be found in textbooks on calculus. Furthermore, the integral is a cornerstone in a systematic theory of probability that allows for the definition and investigation of expected values and higher moments of random variables.

In this chapter, we define the integral by an approximation scheme with simple functions. Then we deduce basic statements such as Fatou's lemma. Other important convergence theorems for integrals follow in Chapters 6 and 7.

4.1 Construction and Simple Properties

In the following, $(\Omega, \mathcal{A}, \mu)$ will always be a measure space. We denote by $\mathbb{E}$ the vector space of simple functions (see Definition 1.93) on $(\Omega, \mathcal{A})$ and by

$$\mathbb{E}^+ := \{f \in \mathbb{E} : f \geq 0\}$$

the cone (why this name?) of nonnegative simple functions. If

$$f = \sum_{i=1}^{m} \alpha_i \mathbb{1}_{A_i} \tag{4.1}$$

for some $m \in \mathbb{N}$ and for $\alpha_1, \ldots, \alpha_m \in (0, \infty)$, and for mutually disjoint sets $A_1, \ldots, A_m \in \mathcal{A}$, then (4.1) is said to be a *normal representation* of f.

Lemma 4.1 *If $f = \sum_{i=1}^m \alpha_i \mathbb{1}_{A_i}$ and $f = \sum_{j=1}^n \beta_j \mathbb{1}_{B_j}$ are two normal representations of $f \in \mathbb{E}^+$, then*

$$\sum_{i=1}^{m} \alpha_i \mu(A_i) = \sum_{j=1}^{n} \beta_j \mu(B_j).$$

A. Klenke, *Probability Theory*, Universitext, DOI 10.1007/978-1-4471-5361-0_4,

Proof If $\mu(A_i \cap B_j) > 0$ for some i and j, then $A_i \cap B_j \neq \emptyset$, and $f(\omega) = \alpha_i = \beta_j$ for any $\omega \in A_i \cap B_j$. Furthermore, clearly $A_i \subset \bigcup_{j=1}^{n} B_j$ if $\alpha_i \neq 0$, and $B_j \subset \bigcup_{i=1}^{m} A_i$ if $\beta_j \neq 0$. We conclude that

$$\begin{aligned}\sum_{i=1}^{m} \alpha_i \mu(A_i) &= \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \mu(A_i \cap B_j) \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} \beta_j \mu(A_i \cap B_j) = \sum_{j=1}^{n} \beta_j \mu(B_j).\end{aligned}$$

□

This lemma allows us to make the following definition (since the value of $I(f)$ does not depend on the choice of the normal representation).

Definition 4.2 Define the map $I : \mathbb{E}^+ \to [0, \infty]$ by

$$I(f) = \sum_{i=1}^{m} \alpha_i \mu(A_i)$$

if f has the normal representation $f = \sum_{i=1}^{m} \alpha_i \mathbb{1}_{A_i}$.

Lemma 4.3 *The map I is positive linear and monotone increasing*: *Let $f, g \in \mathbb{E}^+$ and $\alpha \geq 0$. Then the following statements hold.*

(i) $I(\alpha f) = \alpha I(f)$.
(ii) $I(f + g) = I(f) + I(g)$.
(iii) *If $f \leq g$, then $I(f) \leq I(g)$.*

Proof This is left as an exercise. □

Definition 4.4 (Integral) If $f : \Omega \to [0, \infty]$ is measurable, then we define the *integral* of f with respect to μ by

$$\int f \, d\mu := \sup\big\{I(g) : g \in \mathbb{E}^+, g \leq f\big\}.$$

Remark 4.5 By Lemma 4.3(iii), we have $I(f) = \int f \, d\mu$ for any $f \in \mathbb{E}^+$. Hence the integral is an extension of the map I from $\mathbb{E}^+$ to the set of nonnegative measurable functions. ◊

If $f, g : \Omega \to \overline{\mathbb{R}}$ with $f(\omega) \leq g(\omega)$ for any $\omega \in \Omega$, then we write $f \leq g$. Analogously, we write $f \geq 0$ and so on. On the other hand, we write "$f \leq g$ almost everywhere" if the weaker condition holds that there exists a μ-null set N such that $f(\omega) \leq g(\omega)$ for any $\omega \in N^c$.

Lemma 4.6 *Let $f, g, f_1, f_2, \ldots$ be measurable maps $\Omega \to [0, \infty]$. Then*:

(i) (*Monotonicity*) *If* $f \le g$, *then* $\int f\,d\mu \le \int g\,d\mu$.
(ii) (*Monotone convergence*) *If* $f_n \uparrow f$, *then the integrals also converge*: $\int f_n\,d\mu \uparrow \int f\,d\mu$.
(iii) (*Linearity*) *If* $\alpha, \beta \in [0,\infty]$, *then*

$$\int (\alpha f + \beta g)\,d\mu = \alpha \int f\,d\mu + \beta \int g\,d\mu,$$

where we use the convention $\infty \cdot 0 := 0$.

Proof (i) This is immediate from the definition of the integral.

(ii) By (i), we have

$$\lim_{n\to\infty} \int f_n\,d\mu = \sup_{n\in\mathbb{N}} \int f_n\,d\mu \le \int f\,d\mu.$$

Hence we only have to show $\int f\,d\mu \le \sup_{n\in\mathbb{N}} \int f_n\,d\mu$.

Let $g \in \mathbb{E}^+$ with $g \le f$. It is enough to show that

$$\sup_{n\in\mathbb{N}} \int f_n\,d\mu \ge \int g\,d\mu. \tag{4.2}$$

Assume that the simple function g has the normal representation $g = \sum_{i=1}^N \alpha_i \mathbb{1}_{A_i}$ for some $\alpha_1, \ldots, \alpha_N \in (0,\infty)$ and mutually disjoint sets $A_1, \ldots, A_N \in \mathcal{A}$. For any $\varepsilon > 0$ and $n \in \mathbb{N}$, define the set

$$B_n^\varepsilon = \big\{ f_n \ge (1-\varepsilon) g \big\}.$$

Since $f_n \uparrow f \ge g$, we have $B_n^\varepsilon \uparrow \Omega$ for any $\varepsilon > 0$. Hence, by (i), for any $\varepsilon > 0$,

$$\begin{aligned}\int f_n\,d\mu &\ge \int \big((1-\varepsilon) g \mathbb{1}_{B_n^\varepsilon}\big)\,d\mu \\ &= \sum_{i=1}^N (1-\varepsilon)\alpha_i\, \mu\big(A_i \cap B_n^\varepsilon\big) \stackrel{n\to\infty}{\longrightarrow} \sum_{i=1}^N (1-\varepsilon)\alpha_i\, \mu(A_i) \\ &= (1-\varepsilon) \int g\,d\mu.\end{aligned}$$

Letting $\varepsilon \downarrow 0$ implies (4.2) and hence the claim (ii).

(iii) By Theorem 1.96, any nonnegative measurable map is a monotone limit of simple functions. Hence there are sequences $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ in $\mathbb{E}^+$ such that $f_n \uparrow f$ and $g_n \uparrow g$. Thus also $(\alpha f_n + \beta g_n) \uparrow \alpha f + \beta g$. By (ii) and Lemma 4.3, this implies

$$\int (\alpha f + \beta g)\, d\mu = \lim_{n\to\infty} \int (\alpha f_n + \beta g_n)\, d\mu$$
$$= \alpha \lim_{n\to\infty} \int f_n\, d\mu + \beta \lim_{n\to\infty} \int g_n\, d\mu = \alpha \int f\, d\mu + \beta \int g\, d\mu.$$

□

For any measurable map $f : \Omega \to \overline{\mathbb{R}}$, we have $f^+ \leq |f|$ and $f^- \leq |f|$, which implies $\int f^\pm\, d\mu \leq \int |f|\, d\mu$. In particular, if $\int |f|\, d\mu < \infty$, then also $\int f^-\, d\mu < \infty$ and $\int f^+\, d\mu < \infty$. Thus we can make the following definition that is the final definition for the integral of measurable functions.

Definition 4.7 (Integral of measurable functions) A measurable function $f : \Omega \to \overline{\mathbb{R}}$ is called *μ-integrable* if $\int |f|\, d\mu < \infty$. We write

$$\mathcal{L}^1(\mu) := \mathcal{L}^1(\Omega, \mathcal{A}, \mu) := \left\{ f : \Omega \to \overline{\mathbb{R}} : f \text{ is measurable and } \int |f|\, d\mu < \infty \right\}.$$

For $f \in \mathcal{L}^1(\mu)$, we define the integral of f with respect to μ by

$$\int f(\omega)\, \mu(d\omega) := \int f\, d\mu := \int f^+\, d\mu - \int f^-\, d\mu. \tag{4.3}$$

If we only have $\int f^-\, d\mu < \infty$ or $\int f^+\, d\mu < \infty$, then we also define $\int f\, d\mu$ by (4.3). Here the values $+\infty$ and $-\infty$, respectively, are possible.
For $A \in \mathcal{A}$, we define $\displaystyle\int_A f\, d\mu := \int (f \mathbb{1}_A)\, d\mu$.

Theorem 4.8 *Let $f : \Omega \to [0, \infty]$ be a measurable map.*

(i) *We have $f = 0$ almost everywhere if and only if $\int f\, d\mu = 0$.*
(ii) *If $\int f\, d\mu < \infty$, then $f < \infty$ almost everywhere.*

Proof (i) "$\Longrightarrow$" Assume $f = 0$ almost everywhere. Let $N = \{\omega : f(\omega) > 0\}$. Then $f \leq \infty \cdot \mathbb{1}_N$ and $n\mathbb{1}_N \uparrow \infty \cdot \mathbb{1}_N$. From Lemma 4.6(i) and (ii), we infer

$$0 \leq \int f\, d\mu \leq \int (\infty \cdot \mathbb{1}_N)\, d\mu = \lim_{n\to\infty} \int n\mathbb{1}_N\, d\mu = 0.$$

"$\Longleftarrow$" Let $N_n = \{f \geq \frac{1}{n}\}$, $n \in \mathbb{N}$. Then $N_n \uparrow N$ and

$$0 = \int f\, d\mu \geq \int \frac{1}{n}\mathbb{1}_{N_n}\, d\mu = \frac{\mu(N_n)}{n}.$$

Hence $\mu(N_n) = 0$ for any $n \in \mathbb{N}$ and thus $\mu(N) = 0$.

(ii) Let $A = \{\omega : f(\omega) = \infty\}$. For $n \in \mathbb{N}$, we have $\frac{1}{n} f \mathbb{1}_{\{f \geq n\}} \geq \mathbb{1}_{\{f \geq n\}}$. Hence Lemma 4.6(i) implies

$$\mu(A) = \int \mathbb{1}_A \, d\mu \leq \int \mathbb{1}_{\{f \geq n\}} \, d\mu \leq \frac{1}{n} \int f \mathbb{1}_{\{f \geq n\}} \, d\mu \leq \frac{1}{n} \int f \, d\mu \overset{n\to\infty}{\longrightarrow} 0. \qquad \square$$

Theorem 4.9 (Properties of the integral) *Let $f, g \in \mathcal{L}^1(\mu)$.*

(i) (*Monotonicity*) *If $f \leq g$ almost everywhere, then $\int f \, d\mu \leq \int g \, d\mu$. In particular, if $f = g$ almost everywhere, then $\int f \, d\mu = \int g \, d\mu$.*
(ii) (*Triangle inequality*) $\left| \int f \, d\mu \right| \leq \int |f| \, d\mu$.
(iii) (*Linearity*) *If $\alpha, \beta \in \mathbb{R}$, then $\alpha f + \beta g \in \mathcal{L}^1(\mu)$ and*

$$\int (\alpha f + \beta g) \, d\mu = \alpha \int f \, d\mu + \beta \int g \, d\mu.$$

This equation also holds if at most one of the integrals $\int f \, d\mu$ and $\int g \, d\mu$ is infinite.

Proof (i) Clearly, $f^+ \leq g^+$ and $f^- \geq g^-$ a.e. Hence, by Lemma 4.6(i),

$$\int f^+ \, d\mu \leq \int g^+ \, d\mu \quad \text{and} \quad \int f^- \, d\mu \geq \int g^- \, d\mu.$$

This implies

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu \leq \int g^+ \, d\mu - \int g^- \, d\mu = \int g \, d\mu.$$

(ii) Since $f^+ + f^- = |f|$, Lemma 4.6(iii) yields

$$\left| \int f \, d\mu \right| = \left| \int f^+ \, d\mu - \int f^- \, d\mu \right| \leq \int f^+ \, d\mu + \int f^- \, d\mu$$
$$= \int \left(f^+ + f^-\right) d\mu = \int |f| \, d\mu.$$

(iii) Since $|\alpha f + \beta g| \leq |\alpha| \cdot |f| + |\beta| \cdot |g|$, Lemma 4.6(i) and (iii) yield that $\alpha f + \beta g \in \mathcal{L}^1(\mu)$. In order to show linearity, it is enough to check the following three properties.

(a) $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$.
(b) $\int \alpha f \, d\mu = \alpha \int f \, d\mu$ for $\alpha \geq 0$.
(c) $\int (-f) \, d\mu = -\int f \, d\mu$.

(a) We have $(f + g)^+ - (f + g)^- = f + g = f^+ - f^- + g^+ - g^-$; hence $(f + g)^+ + f^- + g^- = (f + g)^- + f^+ + g^+$. By Lemma 4.6(iii), we infer

$$\int (f + g)^+ \, d\mu + \int f^- \, d\mu + \int g^- \, d\mu = \int (f + g)^- \, d\mu + \int f^+ \, d\mu + \int g^+ \, d\mu.$$

Hence

$$\begin{aligned}\int (f+g)\,d\mu &= \int (f+g)^+\,d\mu - \int (f+g)^-\,d\mu \\ &= \int f^+\,d\mu - \int f^-\,d\mu + \int g^+\,d\mu - \int g^-\,d\mu \\ &= \int f\,d\mu + \int g\,d\mu.\end{aligned}$$

(b) For $\alpha \geq 0$, we have

$$\int \alpha f\,d\mu = \int \alpha f^+\,d\mu - \int \alpha f^-\,d\mu = \alpha \int f^+\,d\mu - \alpha \int f^-\,d\mu = \alpha \int f\,d\mu.$$

(c) We have

$$\begin{aligned}\int (-f)\,d\mu &= \int (-f)^+\,d\mu - \int (-f)^-\,d\mu \\ &= \int f^-\,d\mu - \int f^+\,d\mu = -\int f\,d\mu.\end{aligned}$$

The supplementary statement is simple and is left as an exercise. □

Theorem 4.10 (Image measure) *Let $(\Omega, \mathcal{A})$ and $(\Omega', \mathcal{A}')$ be measurable spaces, let μ be a measure on $(\Omega, \mathcal{A})$ and let $X : \Omega \to \Omega'$ be measurable. Let $\mu' = \mu \circ X^{-1}$ be the image measure of μ under the map X. Assume that $f : \Omega' \to \overline{\mathbb{R}}$ is μ'-integrable. Then $f \circ X \in \mathcal{L}^1(\mu)$ and*

$$\int (f \circ X)\,d\mu = \int f\,d\big(\mu \circ X^{-1}\big).$$

In particular, if X is a random variable on $(\Omega, \mathcal{A}, \mathbf{P})$, then

$$\int f(x)\mathbf{P}[X \in dx] := \int f(x)\mathbf{P}_X[dx] = \int f\,d\mathbf{P}_X = \int f\big(X(\omega)\big)\mathbf{P}[d\omega].$$

Proof This is left as an exercise. □

Example 4.11 (Discrete measure space) Let $(\Omega, \mathcal{A})$ be a discrete measurable space and let $\mu = \sum_{\omega \in \Omega} \alpha_\omega \delta_\omega$ for certain numbers $\alpha_\omega \geq 0$, $\omega \in \Omega$. A map $f : \Omega \to \mathbb{R}$ is integrable if and only if $\sum_{\omega \in \Omega} |f(\omega)|\alpha_\omega < \infty$. In this case,

$$\int f\,d\mu = \sum_{\omega \in \Omega} f(\omega)\alpha_\omega.$$

◊

Definition 4.12 (Lebesgue integral) Let λ be the Lebesgue measure on $\mathbb{R}^n$ and let $f : \mathbb{R}^n \to \mathbb{R}$ be measurable with respect to $\mathcal{B}^*(\mathbb{R}^n)$–$\mathcal{B}(\mathbb{R})$ (here $\mathcal{B}^*(\mathbb{R}^n)$ is the

Lebesgue σ-algebra; see Example 1.71) and λ-integrable. Then we call

$$\int f\,d\lambda$$

the *Lebesgue integral* of f. If $A \in \mathcal{B}(\mathbb{R}^n)$ and $f : \mathbb{R}^n \to \mathbb{R}$ is measurable (or $f : A \to \mathbb{R}$ is $\mathcal{B}^*(\mathbb{R}^n)|_A$–$\mathcal{B}(\mathbb{R})$-measurable and hence $f\mathbb{1}_A$ is $\mathcal{B}^*(\mathbb{R}^n)$–$\mathcal{B}(\mathbb{R})$-measurable), then we write

$$\int_A f\,d\lambda := \int f\mathbb{1}_A\,d\lambda.$$

Definition 4.13 Let μ be a measure on $(\Omega, \mathcal{A})$ and let $f : \Omega \to [0, \infty)$ be a measurable map. Define the measure ν by

$$\nu(A) := \int (\mathbb{1}_A f)\,d\mu \quad \text{for } A \in \mathcal{A}.$$

We say that $f\mu := \nu$ has *density* f with respect to μ.

Remark 4.14 We still have to show that ν is a measure. To this end, we check the conditions of Theorem 1.36. Clearly, $\nu(\emptyset) = 0$. Finite additivity follows from additivity of the integral (Lemma 4.6(iii)). Lower semicontinuity follows from the monotone convergence theorem (Theorem 4.20). ◊

Theorem 4.15 *We have $g \in \mathcal{L}^1(f\mu)$ if and only if $(gf) \in \mathcal{L}^1(\mu)$. In this case,*

$$\int g\,d(f\mu) = \int (gf)\,d\mu.$$

Proof First note that the statement holds for indicator functions. Then, with the usual arguments, extend it step by step first to simple functions, then to nonnegative measurable functions and finally to signed measurable functions. □

Definition 4.16 For measurable $f : \Omega \to \overline{\mathbb{R}}$, define

$$\|f\|_p := \left(\int |f|^p\,d\mu\right)^{1/p}, \quad \text{if } p \in [1, \infty),$$

and

$$\|f\|_\infty := \inf\big\{K \geq 0 : \mu\big(\{|f| > K\}\big) = 0\big\}.$$

Further, for any $p \in [1, \infty]$, define the vector space

$$\mathcal{L}^p(\mu) := \big\{f : \Omega \to \overline{\mathbb{R}} \text{ is measurable and } \|f\|_p < \infty\big\}.$$

Theorem 4.17 *The map $\|\cdot\|_1$ is a seminorm on $\mathcal{L}^1(\mu)$; that is, for all $f, g \in \mathcal{L}^1(\mu)$ and $\alpha \in \mathbb{R}$,*

$$\begin{aligned} \|\alpha f\|_1 &= |\alpha| \cdot \|f\|_1, \\ \|f+g\|_1 &\le \|f\|_1 + \|g\|_1, \\ \|f\|_1 \ge 0 \quad \text{for all } f \quad &\text{and} \quad \|f\|_1 = 0 \quad \text{if } f = 0 \text{ a.e.} \end{aligned} \tag{4.4}$$

Proof The first and the third statements follow from Theorem 4.9(iii) and Theorem 4.8(i). The second statement follows from Theorem 4.9(i) since $|f+g| \le |f| + |g|$; hence

$$\|f+g\|_1 = \int |f+g|\,d\mu \le \int |f|\,d\mu + \int |g|\,d\mu = \|f\|_1 + \|g\|_1. \qquad \square$$

Remark 4.18 In fact, $\|\cdot\|_p$ is a seminorm on $\mathcal{L}^p(\mu)$ for all $p \in [1,\infty]$. Linearity and positivity are obvious, and the triangle inequality is a consequence of Minkowski's inequality, which we will show in Theorem 7.17. ◊

Theorem 4.19 *Let $\mu(\Omega) < \infty$ and $1 \le p' \le p \le \infty$. Then $\mathcal{L}^p(\mu) \subset \mathcal{L}^{p'}(\mu)$ and the canonical inclusion $i : \mathcal{L}^p(\mu) \hookrightarrow \mathcal{L}^{p'}(\mu)$, $f \mapsto f$ is continuous.*

Proof Let $f \in \mathcal{L}^\infty(\mu)$ and $p' \in [1,\infty)$. Then $|f|^{p'} \le \|f\|_\infty^{p'}$ almost everywhere; hence

$$\int |f|^{p'}\,d\mu \le \int \|f\|_\infty^{p'}\,d\mu = \|f\|_\infty^{p'} \cdot \mu(\Omega) < \infty.$$

Thus $\|f-g\|_{p'} \le \mu(\Omega)^{1/p'}\|f-g\|_\infty$ for $f, g \in \mathcal{L}^\infty(\mu)$ and hence i is continuous.

Now let $p, p' \in [1,\infty)$ with $p' < p$ and let $f \in \mathcal{L}^p(\mu)$. Then $|f|^{p'} \le 1 + |f|^p$; hence

$$\int |f|^{p'}\,d\mu \le \mu(\Omega) + \int |f|^p\,d\mu < \infty.$$

Finally, let $f, g \in \mathcal{L}^p(\mu)$. For any $c > 0$, we have

$$|f-g|^{p'} = |f-g|^{p'}\mathbb{1}_{\{|f-g|\le c\}} + |f-g|^{p'}\mathbb{1}_{\{|f-g|>c\}} \le c^{p'} + c^{p'-p}|f-g|^p.$$

In particular, letting $c = \|f-g\|_p$ we obtain

$$\|f-g\|_{p'} \le \left(c^{p'}\mu(\Omega) + c^{p'-p}\|f-g\|_p^p\right)^{1/p'} = \left(1+\mu(\Omega)\right)^{1/p'}\|f-g\|_p.$$

Hence, also in this case, i is continuous. $\square$

Exercise 4.1.1 (Sequence spaces) Now we do not assume $\mu(\Omega) < \infty$. Assume there exists an $a > 0$ such that for any $A \in \mathcal{A}$ either $\mu(A) = 0$ or $\mu(A) \ge a$. Show that the reverse inclusion to Theorem 4.19 holds,

$$\mathcal{L}^{p'}(\mu) \subset \mathcal{L}^p(\mu) \quad \text{if } 1 \le p' \le p \le \infty. \tag{4.5}$$

Exercise 4.1.2 Let $1 \le p' < p \le \infty$ and let μ be σ-finite but not finite. Show that $\mathcal{L}^p(\mu) \setminus \mathcal{L}^{p'}(\mu) \neq \emptyset$.

4.2 Monotone Convergence and Fatou's Lemma

What are the conditions that allow the interchange of limit and integral? In this section, we derive two simple criteria that prepare us for important applications such as the law of large numbers (Chapter 5). More general criteria will be presented in Chapter 6.

Theorem 4.20 (Monotone convergence, Beppo Levi theorem) *Let $f_1, f_2, \ldots \in \mathcal{L}^1(\mu)$ and let $f : \Omega \to \overline{\mathbb{R}}$ be measurable. Assume $f_n \uparrow f$ a.e. for $n \to \infty$. Then*

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

where both sides can equal $+\infty$.

Proof Let $N \subset \Omega$ be a null set such that $f_n(\omega) \uparrow f(\omega)$ for all $\omega \in N^c$. The functions $f'_n := (f_n - f_1)\mathbb{1}_{N^c}$ and $f' := (f - f_1)\mathbb{1}_{N^c}$ are nonnegative and fulfill $f'_n \uparrow f'$. By Lemma 4.6(ii), we have $\int f'_n \, d\mu \stackrel{n\to\infty}{\longrightarrow} \int f' \, d\mu$. Since $f_n = f'_n + f_1$ a.e. and $f = f' + f_1$ a.e., Theorem 4.9(iii) implies

$$\int f_n \, d\mu = \int f_1 \, d\mu + \int f'_n \, d\mu \stackrel{n\to\infty}{\longrightarrow} \int f_1 \, d\mu + \int f' \, d\mu = \int f \, d\mu. \qquad \square$$

Theorem 4.21 (Fatou's lemma) *Let $f \in \mathcal{L}^1(\mu)$ and let $f_1, f_2, \ldots$ be measurable with $f_n \ge f$ a.e. for all $n \in \mathbb{N}$. Then*

$$\int \Big(\liminf_{n\to\infty} f_n \Big) \, d\mu \le \liminf_{n\to\infty} \int f_n \, d\mu.$$

Proof By considering $(f_n - f)_{n\in\mathbb{N}}$, we may assume $f_n \ge 0$ a.e. for all $n \in \mathbb{N}$. Define

$$g_n := \inf_{m \ge n} f_m.$$

Then $g_n \uparrow \liminf_{m\to\infty} f_m$ as $n \to \infty$, and hence by the monotone convergence theorem (Lemma 4.6(ii)) and by monotonicity, $g_n \le f_n$ (thus $\int g_n \, d\mu \le \int f_n \, d\mu$),

$$\int \liminf_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int g_n \, d\mu \le \liminf_{n\to\infty} \int f_n \, d\mu. \qquad \square$$

Example 4.22 (Petersburg game) By a concrete example, we show that in Fatou's lemma the assumption of an integrable minorant is essential. Consider a gamble in a casino where in each round the player's bet either gets doubled or lost. For example, roulette is such a game. If the player bets on "red", she gets the stake back doubled if the ball lands in a red pocket. Otherwise the bet is lost (for the player, not for the casino). There are 37 pockets (in European roulettes), 18 of which are red, 18 are black and one is green (the zero). Hence, by symmetry, the chance of winning should be $p = 18/37 < \frac{1}{2}$. Now assume the gamble is played again and again. We can model this on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ where $\Omega = \{-1, 1\}^{\mathbb{N}}$, $\mathcal{A} = (2^{\{-1,1\}})^{\otimes\mathbb{N}}$ is the σ-algebra generated by the cylinder sets $[\omega_1, \ldots, \omega_n]$ and $\mathbf{P} = ((1-p)\delta_{-1} + p\delta_1)^{\otimes\mathbb{N}}$ is the product measure. Denote by $D_n : \Omega \to \{-1, 1\}$, $\omega \mapsto \omega_n$ the result of the nth game (for $n \in \mathbb{N}$). If in the ith game the player makes a (random) stake of H_i euros, then the cumulative profit after the nth game is

$$S_n = \sum_{i=1}^{n} H_i D_i.$$

Now assume the gambler adopts the following doubling strategy. In the first round, the stake is $H_1 = 1$. If she wins, then she does not bet any money in the subsequent games; that is, $H_n = 0$ for all $n \geq 2$ if $D_1 = 1$. On the other hand, if she loses, then in the second game she doubles the stake; that is, $H_2 = 2$ if $D_1 = -1$. If she wins the second game, she leaves the casino and otherwise doubles the stake again and so on. Hence we can describe the strategy by the formula

$$H_n = \begin{cases} 0, & \text{if there is an } i \in \{1, \ldots, n-1\} \text{ with } D_i = 1, \\ 2^{n-1}, & \text{else.} \end{cases}$$

Note that H_n depends on $D_1, \ldots, D_{n-1}$ only. That is, it is measurable with respect to $\sigma(D_1, \ldots, D_{n-1})$. Clearly, it is a crucial requirement for any strategy that the decision for the next stake depend only on the information available at that time and not depend on the future results of the gamble.

The probability of no win until the nth game is $(1-p)^n$; hence $\mathbf{P}[S_n = 1 - 2^n] = (1-p)^n$ and $\mathbf{P}[S_n = 1] = 1 - (1-p)^n$. Hence we expect an average gain of

$$\int S_n \, d\mathbf{P} = (1-p)^n \big(1 - 2^n\big) + \big(1 - (1-p)^n\big) = 1 - \big(2(1-p)\big)^n \leq 0$$

since $p \leq \frac{1}{2}$ (in the profitable casinos). We define

$$S = \begin{cases} -\infty, & \text{if } -1 = D_1 = D_2 = \ldots, \\ 1, & \text{else.} \end{cases}$$

Then $S_n \xrightarrow{n\to\infty} S$ a.s. but $\lim_{n\to\infty} \int S_n \, d\mathbf{P} < \int S \, d\mathbf{P} = 1$ since $S = 1$ a.s. By Fatou's lemma, this is possible only if there is no integrable minorant for the sequence $(S_n)_{n\in\mathbb{N}}$. If we define $\tilde{S} := \inf\{S_n : n \in \mathbb{N}\}$, then indeed

$$\mathbf{P}\big[\tilde{S} = 1 - 2^{n-1}\big] = \mathbf{P}[D_1 = \ldots = D_{n-1} = -1 \quad \text{and} \quad D_n = 1] = p(1-p)^{n-1}.$$

Hence $\int \tilde{S} \, d\mathbf{P} = \sum_{n=1}^{\infty} (1 - 2^{n-1}) p (1-p)^{n-1} = -\infty$ since $p \leq \frac{1}{2}$. ◊

Exercise 4.2.1 Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and let $f \in \mathcal{L}^1(\mu)$. Show that for any $\varepsilon > 0$, there is an $A \in \mathcal{A}$ with $\mu(A) < \infty$ and $|\int_A f\,d\mu - \int f\,d\mu| < \varepsilon$.

Exercise 4.2.2 Let $f_1, f_2, \ldots \in \mathcal{L}^1(\mu)$ be nonnegative and such that $\lim_{n\to\infty} \int f_n\,d\mu$ exists. Assume there exists a measurable f with $f_n \overset{n\to\infty}{\longrightarrow} f$ μ-almost everywhere. Show that $f \in \mathcal{L}^1(\mu)$ and

$$\lim_{n\to\infty} \int |f_n - f|\,d\mu = \lim_{n\to\infty} \int f_n\,d\mu - \int f\,d\mu.$$

Exercise 4.2.3 Let $f \in \mathcal{L}^1([0,\infty), \lambda)$ be a Lebesgue integrable function on $[0,\infty)$. Show that for λ-almost all $t \in [0,\infty)$ the series $\sum_{n=1}^{\infty} f(nt)$ converges absolutely.

Exercise 4.2.4 Let λ be the Lebesgue measure on $\mathbb{R}$ and let A be a Borel set with $\lambda(A) < \infty$. Show that for any $\varepsilon > 0$, there is a compact set $C \subset A$, a closed set $D \subset \mathbb{R} \setminus A$ and a continuous map $\varphi : \mathbb{R} \to [0,1]$ with $\mathbb{1}_C \le \varphi \le \mathbb{1}_{\mathbb{R}\setminus D}$ and such that $\|\mathbb{1}_A - \varphi\|_1 < \varepsilon$.

Hint: Use the regularity of Lebesgue measure (Remark 1.67).

Exercise 4.2.5 Let λ be the Lebesgue measure on $\mathbb{R}$, $p \in [1,\infty)$ and let $f \in \mathcal{L}^p(\lambda)$. Show that for any $\varepsilon > 0$, there is a continuous function $h : \mathbb{R} \to \mathbb{R}$ such that $\|f - h\|_p < \varepsilon$.

Hint: Use Exercise 4.2.4 to show the assertion first for indicator functions, then for simple functions and finally for general $f \in \mathcal{L}^p(\lambda)$.

Exercise 4.2.6 Let λ be the Lebesgue measure on $\mathbb{R}$, $p \in [1,\infty)$ and let $f \in \mathcal{L}^p(\lambda)$. A map $h : \mathbb{R} \to \mathbb{R}$ is called a *step function* if there exist $n \in \mathbb{N}$ and numbers $t_0 < t_1 < \ldots < t_n$ and $\alpha_1, \ldots, \alpha_n$ such that $h = \sum_{k=1}^{n} \alpha_k \mathbb{1}_{(t_{k-1}, t_k]}$.

Show that for any $\varepsilon > 0$, there exists a step function h such that $\|f - h\|_p < \varepsilon$.

Hint: Use the approximation theorem for measures (Theorem 1.65) with the semiring of left open intervals to show the assertion first for measurable indicator functions. Then use the approximation arguments as in Exercise 4.2.5.

4.3 Lebesgue Integral Versus Riemann Integral

We show that for Riemann integrable functions the Lebesgue integral and the Riemann integral coincide.

Let $I = [a,b] \subset \mathbb{R}$ be an interval and let λ be the Lebesgue measure on I. Further, consider sequences $t = (t^n)_{n\in\mathbb{N}}$ of partitions $t^n = (t_i^n)_{i=0,\ldots,n}$ of I (i.e., $a = t_0^n < t_1^n < \ldots < t_n^n = b$) that get finer and finer. That is,

$$\left|t^n\right| := \max\left\{t_i^n - t_{i-1}^n : i = 1, \ldots, n\right\} \overset{n\to\infty}{\longrightarrow} 0.$$

Assume that for any $n \in \mathbb{N}$, the partition t^{n+1} is a *refinement* of t^n; that is, $\{t_0^n, \ldots, t_n^n\} \subset \{t_0^{n+1}, \ldots, t_{n+1}^{n+1}\}$.

For any function $f : I \to \mathbb{R}$ and any $n \in \mathbb{N}$, define the nth lower sum and upper sum, respectively, by

$$L_n^t(f) := \sum_{i=1}^n \big(t_i^n - t_{i-1}^n\big) \inf f\big(\big[t_{i-1}^n, t_i^n\big)\big),$$

$$U_n^t(f) := \sum_{i=1}^n \big(t_i^n - t_{i-1}^n\big) \sup f\big(\big[t_{i-1}^n, t_i^n\big)\big).$$

A function $f : I \to \mathbb{R}$ is called Riemann integrable if there exists a t such that the limits of the lower sums and upper sums are finite and coincide. In this case, the value of the limit does not depend on the choice of t, and the Riemann integral of f is defined as (see, e.g., [148])

$$\int_a^b f(x)\,dx := \lim_{n\to\infty} L_n^t(f) = \lim_{n\to\infty} U_n^t(f). \tag{4.6}$$

Theorem 4.23 (Riemann integral and Lebesgue integral) *Let $f : I \to \mathbb{R}$ be Riemann integrable on $I = [a, b]$. Then f is Lebesgue integrable on I with integral*

$$\int_I f\,d\lambda = \int_a^b f(x)\,dx.$$

Proof Choose t such that (4.6) holds. By assumption, there is an $n \in \mathbb{N}$ with $|L_n^t(f)| < \infty$ and $|U_n^t(f)| < \infty$. Hence f is bounded. We can thus replace f by $f + \|f\|_\infty$ and hence assume that $f \geq 0$. Define

$$g_n := f(b)\mathbb{1}_{\{b\}} + \sum_{i=1}^n \big(\inf f\big(\big[t_{i-1}^n, t_i^n\big)\big)\big)\mathbb{1}_{[t_{i-1}^n, t_i^n)},$$

$$h_n := f(b)\mathbb{1}_{\{b\}} + \sum_{i=1}^n \big(\sup f\big(\big[t_{i-1}^n, t_i^n\big)\big)\big)\mathbb{1}_{[t_{i-1}^n, t_i^n)}.$$

As t^{n+1} is a refinement of t^n, we have $g_n \leq g_{n+1} \leq h_{n+1} \leq h_n$. Hence there exist g and h with $g_n \uparrow g$ and $h_n \downarrow h$. By construction, we have $g \leq h$ and

$$\begin{aligned}\int_I g\,d\lambda &= \lim_{n\to\infty} \int_I g_n\,d\lambda = \lim_{n\to\infty} L_n^t(f)\\ &= \lim_{n\to\infty} U_n^t(f) = \lim_{n\to\infty} \int_I h_n\,d\lambda = \int_I h\,d\lambda.\end{aligned}$$

Hence $h = g$ λ-a.e. By construction, $g \le f \le h$, and as limits of simple functions, g and h are $\mathcal{B}(I)$–$\mathcal{B}(\mathbb{R})$-measurable. This implies that, for any $\alpha \in \mathbb{R}$, the set

$$\{f \le \alpha\} = \big(\{g \le \alpha\} \cap \{g = h\}\big) \uplus \big(\{f \le \alpha\} \cap \{g \ne h\}\big)$$

is the union of a $\mathcal{B}(I)$-set with a subset of a null set and is hence in $\mathcal{B}(I)^*$ (the Lebesgue completion of $\mathcal{B}(I)$). Hence f is $\mathcal{B}(I)^*$-measurable. By the monotone convergence theorem (Theorem 4.20), we conclude

$$\int_I f \, d\lambda = \lim_{n\to\infty} \int_I g_n \, d\lambda = \int_a^b f(x)\, dx. \qquad \square$$

Example 4.24 Let $f : [0,1] \to \mathbb{R}$, $x \mapsto \mathbb{1}_{\mathbb{Q}}$. Then clearly f is not Riemann integrable since $L_n(f) = 0$ and $U_n(f) = 1$ for all $n \in \mathbb{N}$. On the other hand, f is Lebesgue integrable with integral $\int_{[0,1]} f \, d\lambda = 0$ because $\mathbb{Q} \cap [0,1]$ is a null set. $\Diamond$

Remark 4.25 An improperly Riemann integrable function f on a one-sided open interval $I = (a, b]$ or $I = [0, \infty)$ is not necessarily Lebesgue integrable. Indeed, the improper integral $\int_0^\infty f(x)\, dx := \lim_{n\to\infty} \int_0^n f(x)\, dx$ is defined by a limit procedure that respects the *geometry* of $\mathbb{R}$. The Lebesgue integral does not do that. For example, the function $f : [0,\infty) \to \mathbb{R}$, $x \mapsto \frac{1}{1+x} \sin(x)$ is improperly Riemann integrable but is not Lebesgue integrable since $\int_{[0,\infty)} |f| \, d\lambda = \infty$. $\Diamond$

On the one hand, improperly Riemann integrable functions need not be Lebesgue integrable. On the other hand, there are Lebesgue integrable functions that are not Riemann integrable (such as $\mathbb{1}_{\mathbb{Q}}$). The geometric interpretation is that the Riemann integral respects the geometry of the integration *domain* by being defined via slimmer and slimmer vertical rectangles. On the other hand, the Lebesgue integral respects the geometry of the *range* by being defined via slimmer and slimmer horizontal strips. In particular, the Lebesgue integral does not make any assumption on the geometry of the domain and is thus more universal than the Riemann integral. In order to underline this, we present the following theorem that will also be useful later.

Theorem 4.26 *Let $f : \Omega \to \mathbb{R}$ be measurable and $f \ge 0$ almost everywhere. Then*

$$\sum_{n=1}^\infty \mu\big(\{f \ge n\}\big) \le \int f \, d\mu \le \sum_{n=0}^\infty \mu\big(\{f > n\}\big) \tag{4.7}$$

and

$$\int f \, d\mu = \int_0^\infty \mu\big(\{f \ge t\}\big)\, dt. \tag{4.8}$$

Proof Define $f' = \lfloor f \rfloor$ and $f'' = \lceil f \rceil$. Then $f' \le f \le f''$ and hence $\int f'\,d\mu \le \int f\,d\mu \le \int f''\,d\mu$. Now the first inequality of (4.7) follows from

$$\begin{aligned}\int f'\,d\mu &= \sum_{k=1}^{\infty} \mu\big(\{f' = k\}\big) \cdot k = \sum_{k=1}^{\infty}\sum_{n=1}^{k} \mu\big(\{f' = k\}\big)\\ &= \sum_{n=1}^{\infty}\sum_{k=n}^{\infty} \mu\big(\{f' = k\}\big)\\ &= \sum_{n=1}^{\infty} \mu\big(\{f' \ge n\}\big) = \sum_{n=1}^{\infty} \mu\big(\{f \ge n\}\big).\end{aligned}$$

Similarly, we infer the second inequality in (4.7) from

$$\int f''\,d\mu = \sum_{n=1}^{\infty} \mu\big(\{f'' \ge n\}\big) = \sum_{n=1}^{\infty} \mu\big(\{f > n-1\}\big).$$

If $g(t) := \mu(\{f \ge t\}) = \infty$ for some $t > 0$, then both sides in (4.8) equal ∞. Hence, in the following, assume $g(t) < \infty$ for all $t > 0$.

For $\varepsilon > 0$ and $k \in \mathbb{N}$, define $g^\varepsilon := g \wedge g(\varepsilon)$, $f^\varepsilon := f \mathbb{1}_{\{f \ge \varepsilon\}}$ and $f_k^\varepsilon = 2^k f^\varepsilon$ as well as

$$\alpha_k^\varepsilon := 2^{-k} \sum_{n=1}^{\infty} \mu\big(\{f^\varepsilon \ge n2^{-k}\}\big).$$

Then $\alpha_k^\varepsilon \overset{k\to\infty}{\longrightarrow} \int_0^\infty g^\varepsilon(t)\,dt$. Furthermore, by (4.7) (with f replaced by f_k^ε), we have

$$\begin{aligned}\alpha_k^\varepsilon &= 2^{-k} \sum_{n=1}^{\infty} \mu\big(\{f_k^\varepsilon \ge n\}\big) \le \int f^\varepsilon\,d\mu\\ &\le 2^{-k} \sum_{n=0}^{\infty} \mu\big(\{f_k^\varepsilon > n\}\big) = 2^{-k} \sum_{n=0}^{\infty} \mu\big(\{f^\varepsilon > n2^{-k}\}\big) \le \alpha_k^\varepsilon + 2^{-k} g(\varepsilon).\end{aligned}$$

Since $2^{-k} g(\varepsilon) \overset{k\to\infty}{\longrightarrow} 0$, we get

$$\int_0^\infty g^\varepsilon(t)\,dt = \int f^\varepsilon\,d\mu.$$

Since $f^\varepsilon \uparrow f$ and $g^\varepsilon \uparrow g$ for $\varepsilon \downarrow 0$, the monotone convergence theorem implies (4.8). □

Exercise 4.3.1 Let $f : [0, 1] \to \mathbb{R}$ be bounded. Show that f is (properly) Riemann integrable if and only if f is λ-a.e. continuous.

Exercise 4.3.2 If $f : [0, 1] \to \mathbb{R}$ is Riemann integrable, then f is Lebesgue measurable. Give an example that shows that f need not be Borel measurable.

Hint: Without proof, use the existence of a subset of [0, 1] that is not Borel measurable. Based on this, construct a set that is not Borel and whose closure is a null set.

Exercise 4.3.3 Let $f : [0, 1] \to (0, \infty)$ be Riemann integrable. Without using the equivalence of the Lebesgue integral and the Riemann integral, show that $\int_0^1 f(x)\, dx > 0$.