4 Integral

4a Introduction **34**
4b Simple functions (unsigned) **35**
4c Measurable functions (unsigned) **36**
4d Integrable functions **42**

Lebesgue integral: definition, basic properties. Integral as a new measure. Integral w.r.t. pushforward measure.

4a Introduction

Given a measure space (X, S, μ) and a measurable function $f : X \to [0, \infty]$, we are interested in a measure ν on (X, S) such that

(4a1) $$\mu(A) \inf_{x \in A} f(x) \le \nu(A) \le \mu(A) \sup_{x \in A} f(x) \quad \text{for all } A \in S \,,$$

in order to define the integral by

$$\int_A f \, \mathrm{d}\mu = \nu(A) \,.$$

In symbols, the relation between μ, f and ν is often written as

$$\frac{\mathrm{d}\nu}{\mathrm{d}\mu} = f \,,$$

less often as $\mathrm{d}\nu = f \, \mathrm{d}\mu$, and sometimes[1] as $\nu = f \cdot \mu$; the latter notation is used below.

We start with "simple functions", then proceed to measurable functions $X \to [0, \infty]$ ("unsigned"), and then to measurable functions $X \to [-\infty, \infty]$ ("signed").

Throughout, (X, S, μ) is a measure space.

[1] See for example Def. 6 in Appendix A5 to lecture notes by Klaus Ritter; there, find "Probability theory (WS 2011/12)".

4b Simple functions (unsigned)

4b1 Remark. (a) If μ is a measure and $c \in [0,\infty)$, then $c\mu$ is a measure. (By convention, $0 \cdot \infty = 0$.)

(b) If μ_1, μ_2 are measures, then $\mu_1 + \mu_2$ is a measure. (All measures are on the same (X, S), of course.)

(c) If μ is a measure and $B \in S$, then $A \mapsto \mu(A \cap B)$ is a measure.

By a *simple function*[1] we mean a measurable function $f : X \to \mathbb{R}$ such that $f(X) \subset \mathbb{R}$ is a finite set. For now we assume also $f(X) \subset [0,\infty)$ and call f an unsigned simple function.

4b2 Lemma. For every unsigned simple function f there exists one and only one measure ν satisfying (4a1); and this ν is given by

$$\nu(A) = \sum_{y \in f(X)} y\, \mu\big(A \cap f^{-1}(y)\big) \quad \text{for } A \in S\,.$$

Proof. Uniqueness: it follows from (4a1) that $\nu(A) = y\mu(A)$ whenever $A \subset f^{-1}(y)$; and in general, $\nu(A) = \nu\big(\uplus_y (A \cap f^{-1}(y))\big) = \sum_y \nu\big(A \cap f^{-1}(y)\big) = \sum_y y\mu\big(A \cap f^{-1}(y)\big)$.

Existence: the latter formula gives a measure (by 4b1) and, denoting $b = \sup_{x\in A} f(x)$ we have $A = \uplus_{y \le b}\big(A \cap f^{-1}(y)\big)$ and therefore $\nu(A) \le b \sum_{y\le b} \mu\big(A \cap f^{-1}(y)\big) = b\mu(A)$; the infimum is treated similarly. $\square$

We denote this measure ν by $f \cdot \mu$;

$$(f \cdot \mu)(A) = \sum_{y \in f(X)} y\, \mu\big(A \cap f^{-1}(y)\big)\,.$$

In particular,

$$(\mathbb{1}_B \cdot \mu)(A) = \mu(A \cap B)\,; \tag{4b3}$$

$$(\mathbb{1}_A \cdot \mu)(X) = \mu(A)\,. \tag{4b4}$$

4b5 Exercise.

$$(f \cdot \mu)(A) = \int_0^\infty \mu\big(A \cap f^{-1}(y,\infty)\big)\, \mathrm{d}y\,.$$

(Just the Riemann integral of a step function with bounded support.)

Prove it.

[1] But note that "simple" functions are much more complicated than step functions. Indeed, the indicator of a measurable set is a simple function, even if the set is quite complicated.

Clearly, $(cf) \cdot \mu = c(f \cdot \mu)$ for $c \in [0,\infty)$. Also, if f is constant, $f(\cdot) = c$, then $f \cdot \mu = c\mu$.

4b6 Lemma. $(f+g) \cdot \mu = f \cdot \mu + g \cdot \mu$ for all unsigned simple functions f, g.

Proof. If A is such that f and g are constant on A, then $\big((f+g)\cdot\mu\big)(A) = (f\cdot\mu)(A) + (g\cdot\mu)(A)$ (think, why). And in general, this equality still holds, since A is the disjoint union of such sets:

$$A = \biguplus_{y \in f(X), z \in g(X)} \big(A \cap f^{-1}(y) \cap g^{-1}(z) \big) .$$

□

One says that the map $f \mapsto f \cdot \mu$ is positively linear.

4b7 Exercise. $(fg) \cdot \mu = g \cdot (f \cdot \mu)$ for all unsigned simple functions f, g. Prove it.[1]

In particular,

$$(g \cdot \mu)(A) = \big((g \mathbb{1}_A) \cdot \mu \big)(X) , \tag{4b8}$$

since both sides are equal to $\big(\mathbb{1}_A \cdot (g \cdot \mu) \big)(X)$.

4c Measurable functions (unsigned)

4c1 Definition. The (Lebesgue) integral of a measurable function $f : X \to [0,\infty]$ over a set $A \in S$ is

$$\int_A f \, \mathrm{d}\mu = \sup\{ (g \cdot \mu)(A) : \text{unsigned simple } g \le f \} .$$

Immediate consequences (check them):

$$\text{if } f \text{ is simple, then} \quad \int_A f \, \mathrm{d}\mu = (f \cdot \mu)(A) ; \qquad \text{(simple)} \tag{4c2}$$

$$\text{if } f = g \text{ on } A, \text{ then} \quad \int_A f \, \mathrm{d}\mu = \int_A g \, \mathrm{d}\mu ; \qquad \text{(locality)} \tag{4c3}$$

$$\text{if } f \le g \text{ on } A, \text{ then} \quad \int_A f \, \mathrm{d}\mu \le \int_A g \, \mathrm{d}\mu ; \qquad \text{(monotonicity)} \tag{4c4}$$

$$\text{if } f = c \text{ on } A, \text{ then} \quad \int_A f \, \mathrm{d}\mu = c\mu(A) ; \qquad \text{(constant)} \tag{4c5}$$

---

[1] Hint: similar to 4b6.

(4c6) if $a \le f \le b$ on A, then $a\mu(A) \le \int_A f \, \mathrm{d}\mu \le b\mu(A)$. (mean value)

In probability theory, the (mathematical) expectation of a random variable $X : \Omega \to [0, \infty]$ on a probability space $(\Omega, \mathcal{F}, P)$ is, by definition,

$$\mathbb{E}\, X = \int_\Omega X \, \mathrm{d}P \, .$$

We'll see soon that the map $A \mapsto \int_A f \, \mathrm{d}\mu$ is a measure, and then we'll denote this measure by $f \cdot \mu$. First, additivity.

4c7 Lemma.

$$\int_{A \uplus B} f \, \mathrm{d}\mu = \int_A f \, \mathrm{d}\mu + \int_B f \, \mathrm{d}\mu$$

whenever $A, B \in S$ are disjoint.

Proof. $\int_{A \uplus B} f \, \mathrm{d}\mu = \sup_g (g \cdot \mu)(A \uplus B) = \sup_g \big((g \cdot \mu)(A) + (g \cdot \mu)(B) \big) \le \sup_g (g \cdot \mu)(A) + \sup_g (g \cdot \mu)(B) = \int_A f \, \mathrm{d}\mu + \int_B f \, \mathrm{d}\mu$; we have to prove that $\int_{A \uplus B} f \, \mathrm{d}\mu \ge \int_A f \, \mathrm{d}\mu + \int_B f \, \mathrm{d}\mu$, that is, $\int_{A \uplus B} f \, \mathrm{d}\mu \ge (g_1 \cdot \mu)(A) + (g_2 \cdot \mu)(B)$ for all simple $g_1, g_2 \le f$. We take $g = \max(g_1, g_2)$ (the pointwise maximum); this is also a simple function, and $g \le f$. Thus, $\int_{A \uplus B} f \, \mathrm{d}\mu \ge (g \cdot \mu)(A \uplus B) = (g \cdot \mu)(A) + (g \cdot \mu)(B) \ge (g_1 \cdot \mu)(A) + (g_2 \cdot \mu)(B)$. □

Second, countable additivity.

4c8 Remark. In Definition 2e2 of a measure, the countable additivity may be replaced with the condition

$$A_k \uparrow A \quad \text{implies} \quad \mu(A_k) \uparrow \mu(A) \, .$$

(Think, why is it equivalent.)

4c9 Lemma.

$$A_k \uparrow A \quad \text{implies} \quad \int_{A_k} f \, \mathrm{d}\mu \uparrow \int_A f \, \mathrm{d}\mu$$

for $A, A_1, A_2, \dots \in S$.

Proof. $\int_A f \, \mathrm{d}\mu = \sup_g (g \cdot \mu)(A) = \sup_g \sup_k (g \cdot \mu)(A_k) = \sup_k \sup_g (g \cdot \mu)(A_k) = \sup_k \int_{A_k} f \, \mathrm{d}\mu$. □

Now we introduce the measure $f \cdot \mu$ by

$$(f \cdot \mu)(A) = \int_A f \, \mathrm{d}\mu \quad \text{for } A \in S \, .$$

The notation is consistent due to (4c2).

4c10 Exercise. (a) If μ is finite and f is bounded,[1] then $f \cdot \mu$ is finite;

(b) if μ is σ-finite and f is finite (everywhere), then $f \cdot \mu$ is σ-finite. Prove it.[2]

4c11 Theorem (Monotone Convergence Theorem)**.** Let functions $f, f_1, f_2, \dots : X \to [0,\infty]$ be measurable, and a set $A \in S$. Then

$$f_k \uparrow f \text{ on } A \quad \text{implies} \quad \int_A f_k \, \mathrm{d}\mu \uparrow \int_A f \, \mathrm{d}\mu \, .$$

4c12 Lemma. Let measurable $f_1, f_2, \dots : X \to [0,\infty]$ and $c \in [0,\infty]$ satisfy $f_1 \le f_2 \le \dots$ and $\forall x \in A \;\; \lim_k f_k(x) \ge c$. Then $\lim_k \int_A f_k \, \mathrm{d}\mu \ge c\mu(A)$.

Proof. It is sufficient to prove that $\lim_k \int_A f_k \, \mathrm{d}\mu \ge bp$ whenever $0 \le b < c$ and $0 \le p < \mu(A)$. Given such b and p, we introduce sets $A_k = \{x \in A : f_k(x) \ge b\}$, note that $A_k \uparrow A$ (think, why) and therefore $\mu(A_k) \uparrow \mu(A)$. For k large enough we have $\mu(A_k) \ge p$. The simple function $g = b\mathbb{1}_{A_k}$ satisfies $g \le f_k$, whence $\int_A f_k \ge (g \cdot \mu)(A) = b\mu(A_k) \ge bp$. □

***Proof** of Theorem 4c11.* Clearly, $\lim_k \int_A f_k \, \mathrm{d}\mu$ exists and cannot exceed $\int_A f \, \mathrm{d}\mu$; we have to prove that $\lim_k \int_A f_k \, \mathrm{d}\mu \ge \int_A f \, \mathrm{d}\mu$, that is, $\lim_k \int_A f_k \, \mathrm{d}\mu \ge (g \cdot \mu)(A)$ for arbitrary simple $g \le f$.

We have $(g \cdot \mu)(A) = \sum_{y \in g(X)} y\mu(A_y)$ where $A_y = A \cap g^{-1}(y)$; and, by 4c7, $\int_A f_k \, \mathrm{d}\mu = \sum_{y \in g(X)} \int_{A_y} f_k \, \mathrm{d}\mu$. For each y, on A_y we have $\lim_k f_k = f \ge g = y$; by Lemma 4c12, $\lim_k \int_{A_y} f_k \, \mathrm{d}\mu \ge y\mu(A_y)$. The sum over $y \in g(X)$ completes the proof. □

4c13 Exercise.

$$\int_A f \, \mathrm{d}\mu = \int_0^\infty \mu\big(A \cap f^{-1}(y,\infty]\big) \, \mathrm{d}y \, .$$

Prove it.[3]

(The right-hand side is the Lebesgue integral on $(0,\infty)$ of the function $y \mapsto \mu\big(A \cap f^{-1}(y,\infty]\big)$.)

In particular, let $A = X$, and (X, S) be $\big([0,\infty], \mathcal{B}[0,\infty]\big)$ (μ being an arbitrary measure on this measurable space), and $f = \mathrm{id} : [0,\infty] \to [0,\infty]$. Then

(4c14)
$$\int_{[0,\infty]} \mathrm{id} \, \mathrm{d}\mu = \int_0^\infty \mu\big((y,\infty]\big) \, \mathrm{d}y \quad \text{for all Borel measures } \mu \text{ on } [0,\infty].$$

[1] Not by $+\infty$, of course.

[2] Hint: (a) easy; (b) use (a).

[3] Hint: 4b5; $f_k \uparrow f$; $f_k^{-1}(y,\infty] \uparrow f^{-1}(y,\infty]$; use 4c11 (twice).

Think twice before writing this $\int_{[0,\infty]}$ as $\int_0^\infty$; the points 0 and ∞ may be atoms of the measure μ.

In probability theory, for a random variable $X : \Omega \to [0,\infty]$, $P\bigl(X^{-1}(x,\infty]\bigr)$ is the probability of the event $X > x$, denoted $\mathbb{P}\bigl(X > x\bigr)$, and we get

$$\mathbb{E}\,X = \int_0^\infty \mathbb{P}\bigl(X > x\bigr)\,\mathrm{d}x\,.$$

Positive linearity of the map $f \mapsto f \cdot \mu$ proved in Sect. 4b for simple f will be generalized soon to measurable f. In other words: positive linearity of $\int_A$ (for every given $A \in S$).

For every measurable f there exist simple f_k such that $f_k \uparrow f$. Just choose finite sets $E_1 \subset E_2 \subset \dots \subset [0,\infty)$ whose union is dense in $[0,\infty)$, and take $f_k(x) = \max\{y \in E_k : y \le f(x)\}$.

4c15 Proposition. $\int_A (f+g)\,\mathrm{d}\mu = \int_A f\,\mathrm{d}\mu + \int_A g\,\mathrm{d}\mu$ for all measurable $f, g : X \to [0,\infty]$.

Proof. We take simple f_k, g_k such that $f_k \uparrow f$, $g_k \uparrow g$; then $f_k+g_k \uparrow f+g$. By 4c11, $\int_A f_k\,\mathrm{d}\mu \uparrow \int_A f\,\mathrm{d}\mu$, $\int_A g_k\,\mathrm{d}\mu \uparrow \int_A g\,\mathrm{d}\mu$, and $\int_A (f_k+g_k)\,\mathrm{d}\mu \uparrow \int_A (f+g)\,\mathrm{d}\mu$. Thus, $\int_A (f+g)\,\mathrm{d}\mu = \lim_k \int_A (f_k+g_k)\,\mathrm{d}\mu = \lim_k \bigl(\int_A f_k\,\mathrm{d}\mu + \int_A g_k\,\mathrm{d}\mu\bigr) = \lim_k \int_A f_k\,\mathrm{d}\mu + \lim_k \int_A g_k\,\mathrm{d}\mu = \int_A f\,\mathrm{d}\mu + \int_A g\,\mathrm{d}\mu$. $\square$

Also, $\int_A (cf)\,\mathrm{d}\mu = c\int_A f\,\mathrm{d}\mu$ for $c \ge 0$ (think, why); thus, $\int_A$ is positively linear.

4c16 Corollary (of 4c15 and 4c11)**.** $\int_A \bigl(\sum_{k=1}^\infty f_k\bigr)\,\mathrm{d}\mu = \sum_{k=1}^\infty \int_A f_k\,\mathrm{d}\mu$.

4c17 Exercise. [1,2] Let $f = 0$ on the Cantor set, and $f = k$ on each interval of length 3^{-k} which has been removed from $[0,1]$. Find $\int_{[0,1]} f\,\mathrm{d}m$.

In terms of monotone convergence of measures,

$$\mu_k \uparrow \mu \quad \iff \quad \forall A \in S \;\; \mu_k(A) \uparrow \mu(A)\,, \tag{4c18}$$

the Monotone Convergence Theorem 4c11 becomes

$$f_k \uparrow f \quad \Longrightarrow \quad f_k \cdot \mu \uparrow f \cdot \mu\,; \tag{4c19}$$

and 4c16 becomes

$$(f_1 + f_2 + \dots) \cdot \mu = f_1 \cdot \mu + f_2 \cdot \mu + \dots \tag{4c20}$$

[1] Capiński & Kopp, Exer. 4.2.

[2] Hint: $\sum_{k=1}^\infty k x^{k-1} = \frac{\mathrm{d}}{\mathrm{d}x} \sum_{k=0}^\infty x^k = 1/(1-x)^2$ for $-1 < x < 1$.

4c21 Exercise. Let (Y,T) be a measurable space, $\varphi : X \to Y$ a measurable map, and $f : Y \to [0,\infty]$ a measurable function. Then

$$f \cdot \varphi_* \mu = \varphi_*\big((f \circ \varphi) \cdot \mu\big) \, .$$

Prove it.[1]

We get a "change of variable formula":[2]

$$\int_B f \, \mathrm{d}(\varphi_* \mu) = \int_{\varphi^{-1}(B)} (f \circ \varphi) \, \mathrm{d}\mu \quad \text{for } B \in T \, ; \tag{4c22}$$

$$\int_Y f \, \mathrm{d}(\varphi_* \mu) = \int_X (f \circ \varphi) \, \mathrm{d}\mu \, . \tag{4c23}$$

In particular, let (Y,T) be $\big([0,\infty], \mathcal{B}[0,\infty]\big)$, and $f = \mathrm{id} : [0,\infty] \to [0,\infty]$; we also rename φ to f and get

$$\int_X f \, \mathrm{d}\mu = \int_{[0,\infty]} \mathrm{id} \; \mathrm{d}(f_* \mu) \, ;$$

this fact follows also from 4c13 and (4c14).

In probability theory, for a random variable $X : \Omega \to [0,\infty]$, $X_* P$ is the distribution of X, denoted P_X (as was noted before 3d3), and we get

$$\mathbb{E}\, X = \int_{[0,\infty]} \mathrm{id} \; \mathrm{d}P_X$$

and, more generally, $\mathbb{E}\, f(X) = \int f \, \mathrm{d}P_X$ for Borel $f : [0,\infty] \to [0,\infty]$.

Another special case of 4c21: $Y = X$, $T \subset S$, $\varphi = \mathrm{id}$. In this case $\varphi_* \mu = \mu|_T$; (4c22) becomes

$$\int_B f \, \mathrm{d}(\mu|_T) = \int_B f \, \mathrm{d}\mu$$

for $B \in T$ and T-measurable f. Extending a measure from T to S we do not change integrals that were defined before. In particular, completion of a measure does not change integrals that were defined before the completion.

Extension of the set X may be treated similarly.

4c24 Remark. Every increasing sequence of measures converges to some measure.

Proof (sketch). Let $\mu_i \uparrow \mu$; clearly, μ is additive; countable additivity (similar to 4c9): let $A_j \uparrow A$, then $\mu(A) = \sup_i \mu_i(A) = \sup_i \sup_j \mu_i(A_j) = \sup_j \sup_i \mu_i(A_j) = \sup_j \mu(A_j)$.

[1] Hint: first, f is an indicator; second, f is simple; third, the general case.

[2] Tao, Exer. 1.4.37; Capiński & Kopp Th. 4.41.

4c25 Exercise. $\mu_k \uparrow \mu$ implies $f \cdot \mu_k \uparrow f \cdot \mu$ for unsigned simple f.
Prove it.[1]

4c26 Exercise. $(fg) \cdot \mu = g \cdot (f \cdot \mu)$ for all unsigned measurable f, g.
Prove it.[2]

In particular, if $f : X \to (0, \infty)$, then $\frac{1}{f} \cdot (f \cdot \mu) = \big(\frac{1}{f} f\big) \cdot \mu = 1 \cdot \mu = \mu$; that is,

$$\nu = f \cdot \mu \quad \Longrightarrow \quad \mu = \frac{1}{f} \cdot \nu \quad \text{for } 0 < f < \infty \,.$$

In more traditional notation

$$f = \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \quad \text{for} \quad \nu = f \cdot \mu \tag{4c27}$$

the fact 4c26 becomes

$$\int_A g \, \mathrm{d}\nu = \int_A \Big(g \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \Big) \, \mathrm{d}\mu \,. \tag{4c28}$$

4c29 Example. The standard normal distribution on $\mathbb{R}$ (called also the standard Gaussian measure on $\mathbb{R}$) is the probability measure $\gamma = \varphi \cdot m$ where

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} \quad \text{is the standard normal density} \,.$$

If a random variable X is distributed γ (that is, $P_X = \gamma$), then

$$\mathbb{E} \, f(X) = \int_\Omega f(X) \, \mathrm{d}P = \int_{\mathbb{R}} f \, \mathrm{d}\gamma = \int_{\mathbb{R}} f\varphi \, \mathrm{d}m = \int_{-\infty}^{+\infty} f(t)\varphi(t) \, \mathrm{d}t$$

for every Borel $f : \mathbb{R} \to [0, \infty]$.

4c30 Exercise. (a) $c\mu\{x \in A : f(x) \ge c\} \le \int_A f \, \mathrm{d}\mu$ for all $c \in [0, \infty]$;[3]
(b) if $\int_A f \, \mathrm{d}\mu < \infty$, then $\{x \in A : f(x) = \infty\}$ is a null set;
(c) if $\int_A f \, \mathrm{d}\mu = 0$, then $\{x \in A : f(x) > 0\}$ is a null set.
Prove it.[4]

One says that $f < \infty$ almost everywhere on A, if $\{x \in A : f(x) = \infty\}$ is a sub-null set. (For measurable f it is then a null set.) More generally, given a property of a point of A, one says that this property holds *almost everywhere*

[1]Hint: f is a linear combination of indicators; use (4b3).
[2]Hint: first, do it for simple g using (4b7), $f_k \uparrow f$ and 4c25; second, $g_k \uparrow g$, use (4c19).
[3]Do not forget: $0 \cdot \infty = 0$ (as noted in 4b1).
[4]Hint: (a) integrate f over this set; (b), (c) use (a).

(a.e.) on A, if it holds outside of some sub-null set (and then, necessarily, outside of some null set). In probability theory this is called "almost surely" (a.s.). Thus,

$$\text{(4c31)} \qquad \text{if } \int_A f \, \mathrm{d}\mu < \infty, \text{ then} \qquad f \text{ is finite a.e. on } A\,;$$

$$\text{(4c32)} \qquad \text{if } \int_A f \, \mathrm{d}\mu = 0, \text{ then} \qquad f = 0 \text{ a.e. on } A\,.$$

If $\mu(A) < \infty$ and f is finite a.e. on A, but unbounded, then $\int_A f \, \mathrm{d}\mu$ may converge or diverge. But if $f = 0$ a.e. on A, then $\int_A f \, \mathrm{d}\mu = 0$ (even if $\mu(A) = \infty$), since this is evidently true for simple functions. In particular, $\int_Z f \, \mathrm{d}\mu = 0$ for all f, if Z is a null set. (Indeed, even the equality $0 = \infty$ holds a.e. on a null set!) It follows by 4c7 that $\int_A f \, \mathrm{d}\mu = \int_{A \setminus Z} f \, \mathrm{d}\mu$; null sets are negligible.

Two functions are called *equivalent,* if they are equal almost everywhere.

Denoting by $[f]$ the equivalence class of f we may write the equivalence as $[f] = [g]$. If $[f] = [g]$ then $\int_A f \, \mathrm{d}\mu = \int_A g \, \mathrm{d}\mu$ for all A (just because null sets are negligible). That is, $\int_A [f] \, \mathrm{d}\mu$ is well-defined. Also, $[f] \cdot \mu$ is well-defined.

If $[f_1] = [g_1]$ and $[f_2] = [g_2]$, then $[f_1 + f_2] = [g_1 + g_2]$ (think, why); thus, the sum of two equivalence classes is a well-defined equivalence class. Moreover, the same holds for the sum of countably many equivalence classes. Also the relation $[f] \le [g]$ is well-defined.

Functions may be replaced with equivalence classes in all our statements. For instance, in (4c6):

$$\text{if } a \le f \le b \text{ a.e. on } A, \text{ then} \quad a\mu(A) \le \int_A f \, \mathrm{d}\mu \le b\mu(A)\,;$$

in 4c11:

$$f_k \uparrow f \text{ a.e. on } A \quad \text{implies} \quad \int_A f_k \, \mathrm{d}\mu \uparrow \int_A f \, \mathrm{d}\mu\,;$$

and so on. Usually one still writes functions (just for convenience), but means their equivalence classes.

4d Integrable functions

4d1 Definition. A measurable function $f : X \to [-\infty, +\infty]$ is *integrable,* if $\int_X |f| \, \mathrm{d}\mu < \infty$.

Clearly, integrable functions are a vector space. The functional $f \mapsto \int_X |f| \, \mathrm{d}\mu$ is (generally) not a norm on this space of functions, but is a norm on the corresponding space of equivalence classes:

$$\|[f]\| = \int_X |f| \,\mathrm{d}\mu \,;$$
$$\|[cf]\| = |c| \|[f]\| \,;$$
$$\|[f+g]\| \le \|[f]\| + \|[g]\| \,;$$
$$\|[f]\| = 0 \quad \Longleftrightarrow \quad [f] = [0] \,.$$

This normed[1] space is denoted by $L_1(X, S, \mu)$, or just $L_1(\mu)$.[2]

Integrable functions are finite a.e.; WLOG we may assume that they are finite everywhere.

Every integrable function can be written as the difference of two unsigned integrable functions; in particular,

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

4d2 Lemma. If unsigned integrable f_1, f_2, g_1, g_2 satisfy $f_1 - f_2 = g_1 - g_2$, then $\int_X f_1 \,\mathrm{d}\mu - \int_X f_2 \,\mathrm{d}\mu = \int_X g_1 \,\mathrm{d}\mu - \int_X g_2 \,\mathrm{d}\mu$.

Proof. $f_1 + g_2 = f_2 + g_1$; by 4c15, $\int f_1 + \int g_2 = \int f_2 + \int g_1$, that is, $\int f_1 - \int f_2 = \int g_1 - \int g_2$. □

Thus, we may define

$$\int_X f \,\mathrm{d}\mu = \int_X g \,\mathrm{d}\mu - \int_X h \,\mathrm{d}\mu \quad \text{whenever } f = g - h \,;$$

here f is integrable, and g, h are unsigned integrable. Clearly,

$$[f] \mapsto \int_X f \,\mathrm{d}\mu \quad \text{is a linear functional on } L_1(\mu) \,,$$
$$\left| \int_X f \,\mathrm{d}\mu \right| \le \|[f]\| \,.$$

The same holds for $\int_A$, of course.

A vector-function $f : X \to \mathbb{R}^n$, $f(x) = \big(f_1(x), \dots, f_n(x)\big)$, is called integrable, if its coordinate functions $f_1, \dots, f_n$ are integrable; in this case, by definition,

$$\int_A f \,\mathrm{d}\mu = \left(\int_A f_1 \,\mathrm{d}\mu, \dots, \int_A f_n \,\mathrm{d}\mu \right) .$$

[1] In fact, Banach space; its completeness will be proved later.

[2] Or $L^1(\mu)$.

With respect to integrability, complex-valued functions $X \to \mathbb{C}$ may be treated as just $X \to \mathbb{R}^2$ (and $X \to \mathbb{C}^n$ as $X \to \mathbb{R}^{2n}$).

Applying 4c13 to f^+ and f^- we get (for integrable f)

$$\text{(4d3)} \qquad \int_A f \, \mathrm{d}\mu = \int_0^\infty \mu\big(A \cap f^{-1}(y,\infty)\big) - \int_0^\infty \mu\big(A \cap f^{-1}(-\infty,y)\big) \, \mathrm{d}y \, .$$

Similarly to (4c14),

$$\text{(4d4)} \qquad \int_{\mathbb{R}} \mathrm{id} \; \mathrm{d}\mu = \int_0^\infty \mu\big((y,\infty)\big) \, \mathrm{d}y - \int_0^\infty \mu\big((-\infty,-y)\big) \, \mathrm{d}y$$

for all Borel measures μ on $\mathbb{R}$ such that $\int_{\mathbb{R}} |\cdot| \, \mathrm{d}\mu < \infty$.

In probability theory, for an integrable random variable X,

$$\mathbb{E}\, X = \int_0^\infty \mathbb{P}\big(X > x\big) \, \mathrm{d}x - \int_0^\infty \mathbb{P}\big(X < -x\big) \, \mathrm{d}x \, .$$

Applying (4c22) and (4c23) to f^+ and f^- we see that they hold for all integrable f. In particular,

$$\int_X f \, \mathrm{d}\mu = \int_{\mathbb{R}} \mathrm{id} \; \mathrm{d}(f_*\mu) \, ;$$

this fact follows also from 4d3 and (4d4). In probability theory,

$$\mathbb{E}\, X = \int_{\mathbb{R}} \mathrm{id} \; \mathrm{d}P_X \quad \text{for all integrable } X \, ,$$
$$\mathbb{E}\, f(X) = \int_{\mathbb{R}} f \, \mathrm{d}P_X \quad \text{for all } P_X\text{-integrable } f \, .$$

For vector-functions $f : X \to \mathbb{R}^n$, similarly,

$$\int_X f \, \mathrm{d}\mu = \int_{\mathbb{R}^n} \mathrm{id} \; \mathrm{d}(f_*\mu) \, ,$$

μ-integrability of f being equivalent to $(f_*\mu)$-integrability of id. In probability theory,

$$\mathbb{E}\, f(X_1, \dots, X_n) = \int_{\mathbb{R}^n} f \, \mathrm{d}P_{X_1,\dots,X_n} \, ,$$

where $P_{X_1,\dots,X_n} = X_* P$ is the joint distribution (recall the paragraph before 3d3).

Index

a.e., 42
a.s., 42
almost everywhere, 41
almost surely, 42

change of variable, 40

equivalence class, 42
equivalent functions, 42
expectation, 37

integrable function, 42
integrable vector-function, 43
integral, 36, 43

monotone convergence theorem, 38

norm, 43

positively linear, 36

simple function, 35

unsigned, 35

$\frac{d\nu}{d\mu}$, 41
$\mathbb{E}\,X$, 37, 44
$\mathbb{E}\,f(X)$, 40, 44
$[f]$, 42
$|||[f]|||$, 43
$f \cdot \mu$, 35, 37
f^-, f^+, 43
$\int f\,d\mu$, 36, 43
$L_1(\dots)$, 43
$0 \cdot \infty = 0$, 35
P_X, 40