11 Measure theory and integration

Topology does not depend on the notion of 'size'. We do not need to know the length, area or volume of subsets of a given set to understand the topological structure. *Measure theory* is that area of mathematics concerned with the attribution of precisely these sorts of properties. The structure that tells us which subsets are *measurable* is called a *measure space*. It is somewhat analogous with a topological structure, telling us which sets are *open*, and indeed there is a certain amount of interaction between measure theory and topology. A measure space requires firstly an algebraic structure known as a *σ-algebra* imposed on the power set of the underlying space. A *measure* is a positive-valued real function on the σ-algebra that is *countably additive*, whereby the measure of a union of disjoint measurable sets is the sum of their measures. The measure of a set may well be zero or infinite. Full introductions to this subject are given in [1–5], while the flavour of the subject can be found in [6–8].

It is important that measure be not just finitely additive, else it is not far-reaching enough, yet to allow it to be additive on arbitrary unions of disjoint sets would lead to certain contradictions – either all sets would have to be assigned zero measure, or the measure of a set would not be well-defined. By general reckoning the broadest useful measure on the real line or its cartesian products is that due to Lebesgue (1875–1941), and Lebesgue's theory of *integration* based on this theory is in most ways the best definition of integration available.

Use will frequently be made in this chapter of the **extended real line** $\overline{\mathbb{R}}$ consisting of $\mathbb{R} \cup \{\infty\} \cup \{-\infty\}$, having rules of addition $a + \infty = \infty$, $a + (-\infty) = -\infty$ for all $a \in \mathbb{R}$, but no value is given to $\infty + (-\infty)$. The natural order on the real line is supplemented by the inequalities $-\infty < a < \infty$ for all real numbers a. Multiplication can also be extended in some cases, such as $a\infty = \infty$ if $a > 0$, but it is best to avoid the product 0∞ unless a clear convention can be adopted. The natural order topology on $\mathbb{R}$, generated by open intervals (a, b) is readily extended to $\overline{\mathbb{R}}$.

Exercise: Show $\overline{\mathbb{R}}$ is a compact topological space with respect to the order topology.

11.1 Measurable spaces and functions

Measurable spaces

Given a set X, a **σ-algebra** $\mathcal{M}$ on X consists of a collection of subsets, known as **measurable sets**, satisfying

(Meas1) The empty set $\emptyset$ is a measurable set, $\emptyset \in \mathcal{M}$.
(Meas2) If A is measurable then so is its complement:

$$A \in \mathcal{M} \Longrightarrow A^c = X - A \in \mathcal{M}.$$

(Meas3) $\mathcal{M}$ is closed under countable unions:

$$A_1, A_2, A_3, \ldots \in \mathcal{M} \Longrightarrow \bigcup_i A_i \in \mathcal{M}.$$

The pair $(X, \mathcal{M})$ is known as a **measurable space**. Although there are similarities between these axioms and (Top1)–(Top3) for a topological space, (Meas2) is distinctly different in that the complement of an open set is a closed set and is rarely open. It follows from (Meas1) and (Meas2) that the whole space $X = \emptyset^c$ is measurable. The intersection of any pair of measurable sets is measurable, for

$$A \cap B = (A^c \cup B^c)^c.$$

Also, $\mathcal{M}$ is closed with respect to taking differences,

$$A - B = A \cap B^c = (A^c \cup B)^c.$$

Exercise: Show that any countable intersection of measurable sets is measurable.

Example 11.1 Given any set X, the collection $\mathcal{M} = \{\emptyset, X\}$ is obviously a σ-algebra. This is the smallest σ-algebra possible. By contrast, the largest σ-algebra is the set of all subsets 2^X. All interesting examples fall somewhere between these two extremes.

It is trivial to see that the intersection of any two σ-algebras $\mathcal{M} \cap \mathcal{M}'$ is another σ-algebra – check that properties (Meas1)–(Meas3) are satisfied by the sets common to the two σ-algebras. This statement extends to the intersection of an arbitrary family of σ-algebras, $\bigcap_{i \in I} \mathcal{M}_i$. Hence, given any collection of subsets $\mathcal{A} \subset 2^X$, there is a unique 'smallest' σ-algebra $\mathcal{S} \supseteq \mathcal{A}$. This is the intersection of all σ-algebras that contain $\mathcal{A}$. It is called the **σ-algebra generated by** $\mathcal{A}$. For a topological space X, the σ-algebra generated by the open sets are called **Borel sets** on X. They include all open and all closed sets and, in general, many more that are neither open nor closed.

Example 11.2 In the standard topology on the real line $\mathbb{R}$ every open set is a countable union of open intervals. Hence the Borel sets are generated by the set of all open intervals $\{(a, b) \mid a < b\}$. Infinite left-open intervals such as $(a, \infty) = (a, a + 1) \cup (a, a + 2) \cup (a, a + 3) \cup \ldots$ are Borel sets by (Meas3), and similarly all intervals $(-\infty, a)$ are Borel. The complements of these sets are the infinite right or left-closed intervals $(-\infty, a]$ and $[a, \infty)$. Hence all closed intervals $[a, b] = (-\infty, b] \cap [a, \infty)$ are Borel sets.

Exercise: Prove that the σ-algebra of Borel sets on $\mathbb{R}$ is generated by (a) the infinite left-open intervals (a, ∞), (b) the closed intervals $[a, b]$.

Exercise: Prove that all singletons $\{a\}$ are Borel sets on $\mathbb{R}$.

If $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ are two measurable spaces then we define the **product measurable space** $(X \times Y, \mathcal{M} \otimes \mathcal{N})$, by setting the σ-algebra $\mathcal{M} \otimes \mathcal{N}$ on $X \times Y$ to be the σ-algebra generated by all sets of the form $A \times B$ where $A \in \mathcal{M}$ and $B \in \mathcal{N}$.

Measurable functions

Given two measurable spaces $(X, \mathcal{M})$ and $(Y, \mathcal{N})$, a map $f : X \to Y$ is said to be a **measurable function** if the inverse image of every measurable set is measurable:

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

This definition mirrors that for a continuous function in topological spaces.

Theorem 11.1 *If X and Y are topological spaces and $\mathcal{M}$ and $\mathcal{N}$ are the σ-algebras of Borel sets, then every continuous function $f : X \to Y$ is Borel measurable.*

Proof: Let $\mathcal{O}_X$ and $\mathcal{O}_Y$ be the families of open sets in X and Y respectively. We adopt the notation $f^{-1}(\mathcal{A}) \equiv \{f^{-1}(A) \mid A \in \mathcal{A}\}$ for any family of sets $\mathcal{A} \subseteq 2^Y$. Since f is continuous, $f^{-1}(\mathcal{O}_Y) \subseteq \mathcal{O}_X$. The σ-algebras of Borel sets on the two spaces are $\mathcal{M} = \mathcal{S}(\mathcal{O}_X)$ and $\mathcal{N} = \mathcal{S}(\mathcal{O}_Y)$. To prove f is Borel measurable we must show that $f^{-1}(\mathcal{N}) \subseteq \mathcal{M}$. Let

$$\mathcal{N}' = \{B \subseteq Y \mid f^{-1}(B) \in \mathcal{M}\} \subset 2^Y.$$

This is a σ-algebra on Y, for $f^{-1}(\emptyset) = \emptyset$ and

$$f^{-1}(B^c) = \big(f^{-1}(B)\big)^c, \qquad f^{-1}\Big(\bigcup_i B_i\Big) = \bigcup_i f^{-1}(B_i).$$

Hence $\mathcal{N}' \supseteq \mathcal{O}_Y$ for $f^{-1}(\mathcal{O}_Y) \subseteq \mathcal{O}_X \subset \mathcal{M}$. Since $\mathcal{N}$ is the σ-algebra generated by $\mathcal{O}_Y$ we must have that $\mathcal{N}' \supseteq \mathcal{N}$. Hence $f^{-1}(\mathcal{N}) \subseteq f^{-1}(\mathcal{N}') \subseteq \mathcal{M}$ as required. ■

Exercise: If $f : X \to Y$ and $g : Y \to Z$ are measurable functions between measure spaces, show that the composition $g \circ f : X \to Z$ is a measurable function.

If $f : X \to \mathbb{R}$ is a measurable real-valued function on a measurable space $(X, \mathcal{M})$, where $\mathbb{R}$ is assumed given the Borel structure of Example 11.2, it follows that the set

$$\{x \mid f(x) > a\} = f^{-1}\big((a, \infty)\big)$$

is measurable in X. Since the family of Borel sets $\mathcal{B}$ on the real line is generated by the intervals (a, ∞) (see the exercise following Example 11.2), this can actually be used as a criterion for measurability: $f : X \to \mathbb{R}$ is a measurable function iff for any $a \in \mathbb{R}$ the set $\{x \mid f(x) > a\}$ is measurable.

Exercise: Prove the sufficiency of this condition [refer to the proof of Theorem 11.1].

Example 11.3 If $(X, \mathcal{M})$ is a measurable space then the characteristic function $\chi_A : X \to \mathbb{R}$ of a set $A \subset X$ is measurable if and only if $A \in \mathcal{M}$, since for any $a \in \mathbb{R}$

$$\{x \mid \chi_A(x) > a\} = \begin{cases} X & \text{if } a < 0, \\ A & \text{if } 0 \le a < 1, \\ \emptyset & \text{if } a \ge 1. \end{cases}$$

Exercise: Show that for any $a \in R$, the set $\{x \in X \mid f(x) = a\}$ is a measurable set of X.

If $f : X \to \mathbb{R}$ is a measurable function then so is its modulus $|f|$, since the continuous function $x \mapsto |x|$ on $\mathbb{R}$ is necessarily Borel measurable, and $|f|$ is the composition function $|\cdot| \circ f$. Similarly the function f^a for $a > 0$ is measurable, and $1/f$ is measurable if $f(x) \neq 0$ for all $x \in X$. If $g : X \to R$ is another measurable function then the function $f + g$ is measurable since it can be written as the composition $f = \rho \circ F$ where $F : X \to \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ and $\rho : \mathbb{R}^2 \to \mathbb{R}$ are the maps

$$F : x \mapsto (f(x), g(x)) \quad \text{and} \quad \rho(a, b) = a + b.$$

The function F is measurable since the inverse image of any product of intervals I_1, I_2 is

$$F^{-1}(I_1 \times I_2) = f^{-1}(I_1) \cap g^{-1}(I_2),$$

which is a measurable set in X since f and g are assumed measurable functions, while the map ρ is evidently continuous on $\mathbb{R}^2$.

Exercise: Show that for measurable functions $f, g : X \to \mathbb{R}$, the function fg is measurable.

An important class of functions are the **simple functions**: measurable functions $h : X \to \overline{\mathbb{R}}$ that take on only a finite set of extended real values $a_1, a_2, \dots, a_n$. Since $A_i = h^{-1}(\{a_i\})$ is a measurable subset of X for each a_i, we can write a simple function as a linear combination of measurable characteristic functions

$$h = a_1 \chi_{A_1} + a_2 \chi_{A_2} + \cdots + a_n \chi_{A_n} \qquad \text{where all } a_i \neq 0.$$

Some authors use the term *step function* instead of simple function, but the common convention is to preserve this term for simple functions $h : \mathbb{R} \to \mathbb{R}$ in which each set A_i is a union of disjoint *intervals* (Fig. 11.1).

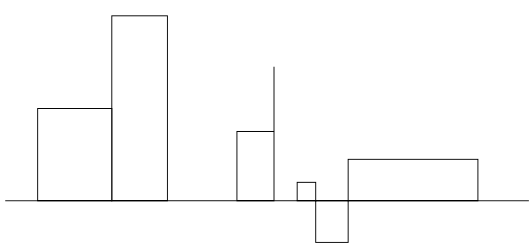

Figure 11.1 Simple (step) function

Let f and g be any pair of measurable functions from X into the extended reals $\overline{R}$. The function $h = \sup(f, g)$ defined by

$$h(x) = \begin{cases} f(x) & \text{if } f(x) \geq g(x) \\ g(x) & \text{if } g(x) > f(x) \end{cases}$$

is measurable, since

$$\{x \mid h(x) > a\} = \{x \mid f(x) > a \text{ or } g(x) > a\} = \{x \mid f(x) > a\} \cup \{x \mid g(x) > a\}$$

is measurable. Similarly, $\inf(f, g)$ is a measurable function. In particular, if f is a measurable function, its positive and negative parts

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

are measurable functions.

Exercise: Show that $f = f^+ - f^-$ and $|f| = f^+ + f^-$.

A simple extension of the above argument shows that $\sup f_n$ is a measurable function for any countable set of measurable functions $f_1, f_2, f_3, \dots$ with values in the extended real numbers $\overline{\mathbb{R}}$. We define the lim sup as

$$\limsup f_n(x) = \inf_{n \geq 1} F_n(x) \quad \text{where} \quad F_n(x) = \sup_{k \geq n} f_k(x).$$

The lim sup always exists since the functions F_n are everywhere monotone decreasing,

$$F_1(x) \geq F_2(x) \geq F_3(x) \geq \dots$$

and therefore have a limit if they are bounded below or approach $-\infty$ if unbounded below. Similarly we can define

$$\liminf f_n(x) = \sup_{n \geq 1} G_n(x) \quad \text{where} \quad G_n(x) = \inf_{k \geq n} f_k(x).$$

It follows that if f_n is a sequence of measurable functions then $\limsup f_n$ and $\liminf f_n$ are also measurable. By standard arguments in analysis $f_n(x)$ is a convergent sequence if and only if $\limsup f_n(x) = \liminf f_n(x) = \lim f_n(x)$. Hence the limit of any convergent sequence of measurable functions $f_n(x) \to f(x)$ is measurable. Note that the convergence need only be 'pointwise convergence', not uniform convergence as is required in many theorems in Riemann integration.

Theorem 11.2 *Any measurable function $f : X \to \mathbb{R}$ is the limit of a sequence of simple functions. The sequence can be chosen to be monotone increasing at all positive values of f, and monotone decreasing at negative values.*

Proof: Suppose f is positive and bounded above, $0 \leq f(x) \leq M$. For each integer $n = 1, 2, \dots$ let h_n be the simple function

$$h_n = \sum_{k=0}^{n} \frac{kM}{2^n} \chi_{A_k}$$

where A_k is the measurable set

$$A_k = \left\{x \;\middle|\; \frac{kM}{2^n} \le f(x) < \frac{k+1}{2^n}\right\}.$$

These simple functions are increasing, $0 \le h_1(x) \le h_2(x) \le \dots$ and $|f(x) - h_n(x)| < M/2^n$. Hence $h_n(x) \to f(x)$ for all $x \in X$ as $n \to \infty$.

If f is any positive function, possibly unbounded, the functions $g_n = \inf(h_n, n)$ are positive and bounded above. Hence for each n there exists a simple function h'_n such that $|g_n - h'_n| < 1/n$. The sequence of simple functions h'_n clearly converges everywhere to $f(x)$. To obtain a monotone increasing sequence of simple functions that converge to f, set

$$f_n = \sup(h'_1, h'_2, \dots, h'_n).$$

If f is not positive, construct simple function sequences approaching the positive and negative parts and use $f = f^+ - f^-$. ■

Problems

Problem 11.1 If $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ are measurable spaces, show that the projection maps $\mathrm{pr}_1 : X \times Y \to X$ and $\mathrm{pr}_2 : X \times Y \to Y$ defined by $\mathrm{pr}_1(x, y) = x$ and $\mathrm{pr}_2(x, y) = y$ are measurable functions.

Problem 11.2 Find a step function $s(x)$ that approximates $f(x) = x^2$ uniformly to within $\varepsilon > 0$ on $[0, 1]$, in the sense that $|f(x) - s(x)| < \varepsilon$ everywhere in $[0, 1]$.

Problem 11.3 Let $f : X \to \mathbb{R}$ and $g : X \to \mathbb{R}$ be measurable functions and $E \subset X$ a measurable set. Show that

$$h(x) = \begin{cases} f(x) & \text{if } x \in E \\ g(x) & \text{if } x \notin E \end{cases}$$

is a measurable function on X.

Problem 11.4 If $f, g : \mathbb{R} \to \mathbb{R}$ are Borel measurable real functions show that $h(x, y) = f(x)g(y)$ is a measurable function $h : \mathbb{R}^2 \to \mathbb{R}$ with respect to the product measure on $\mathbb{R}^2$.

11.2 Measure spaces

Given a measurable space $(X, \mathcal{M})$, a **measure** μ on X is a function $\mu : \mathcal{M} \to \overline{\mathbb{R}}$ such that

(Meas4) $\mu(A) \ge 0$ for all measurable sets A, and $\mu(\emptyset) = 0$.

(Meas5) If $A_1, A_2, \dots \in \mathcal{M}$ is any mutually disjoint sequence (finite or countably infinite) of measurable sets such that $A_i \cap A_j = \emptyset$ if $i \ne j$, then

$$\mu(A_1 \cup A_2 \cup A_3 \cup \dots) = \sum_i \mu(A_i).$$

A function μ satisfying property (Meas5) is often referred to as being *countably additive*. If the series on the right-hand side is not convergent, it is given the value ∞. A **measure space** is a triple $(X, \mathcal{M}, \mu)$ consisting of a measurable space $(X, \mathcal{M})$ together with a measure μ.

Exercise: Show that if $B \subset A$ are measurable sets, then $\mu B \le \mu A$.

Example 11.4 An occasionally useful measure on a σ-algebra $\mathcal{M}$ defined on a set X is the *Dirac measure*. Let a be any fixed point of X, and set

$$\delta_a(A) = \begin{cases} 1 & \text{if } a \in A, \\ 0 & \text{if } a \notin A. \end{cases}$$

(Meas4) holds trivially for $\mu = \delta_a$, and (Meas5) follows from the obvious fact that the union of a disjoint family of sets $\{A_i\}$ can contain a if and only if $a \in A_j$ for precisely one member A_j. This measure has applications in distribution theory, Chapter 12.

Example 11.5 The branch of mathematics known as *probability theory* is best expressed in terms of measure theory. A **probability space** is a measure space $(\Omega, \mathcal{M}, P)$, where $P(\Omega) = 1$. Sets $A \in \Omega$ are known as **events** and Ω is sometimes referred to as the *universe*. This 'universe' is usually thought of as the set of all possible outcomes of a specific experiment. Note that events are not outcomes of the experiment, but *sets* of possible outcomes. The measure function P is known as the **probability measure** on Ω, and $P(A)$ is the **probability of the event** A. All events have probability in the range $0 \le P(A) \le 1$. The element $\emptyset$ has probability 0, $P(\emptyset) = 0$, and is called the **impossible event**. The entire space Ω has probability 1; it can be thought of as the **certainty event**.

The event $A \cup B$ is referred to as **either** A **or** B, while $A \cap B$ is A **and** B. Since P is additive on disjoint sets and $A \cup B = (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$, the probabilities are related by

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

The two events are said to be **independent** if $P(A \cap B) = P(A)P(B)$. This is by no means always the case.

We think of the probability of event B after knowing that A has occurred as the **conditional probability** $P(B|A)$, defined by

$$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

Events A and B are independent if and only if $P(B|A) = P(B)$ – in other words, the probability of B in no way depends on the occurrence of A.

For a finite or countably infinite set of disjoint events $H_1, H_2, \dots$ partitioning Ω (sometimes called a **hypothesis**), $\Omega = \bigcup_i H_i$, we have for any event B

$$B = \bigcup_{i=1}^{\infty} (H_i \cap B).$$

Since the sets in this countable union are mutually disjoint, the probability of B is

$$P(B) = \sum_{i=1}^{\infty} P(H_i \cap B) = \sum_{i=1}^{\infty} P(B|H_i)P(H_i).$$

This leads to **Bayes' formula** for the conditional probability of the hypothesis H_i given the outcome event B,

$$\begin{aligned} P(H_i|B) &= \frac{P(B\cap H_i)}{P(B)} \\ &= \frac{P(H_i)P(B|H_i)}{P(B)} \\ &= \frac{P(H_i)P(B|H_i)}{\sum_{k=1}^{\infty} P(B|H_k)P(H_k)}. \end{aligned}$$

Theorem 11.3 *Let $E_1, E_2, \dots$ be a sequence of measurable sets, which is increasing in the sense that $E_n \subset E_{n+1}$ for all $n = 1, 2, \dots$ Then*

$$E = \bigcup_{n=1}^{\infty} E_n \in \mathcal{M} \quad \textit{and} \quad \mu E = \lim_{n\to\infty} \mu E_n.$$

Proof: E is measurable by condition (Meas3). Set

$$F_1 = E_1,\ F_2 = E_2 - E_1, \dots,\ F_n = E_n - E_{n-1}, \dots.$$

The sets F_n are all measurable and disjoint, $F_n \cap F_m = \emptyset$ if $n \neq m$. Since $E_n = F_1 \cup F_2 \cup \dots \cup F_n$ we have by (Meas5)

$$\lim_{n\to\infty} \mu E_n = \sum_{i=1}^{\infty} \mu(F_n) = \mu\Big(\bigcup_{k=1}^{\infty} F_k\Big) = \mu(E).$$ ■

Lebesgue measure

Every open set U on the real line is a countable union of disjoint open intervals. This follows from the fact that every rational point $r \in U$ lies in a maximal open interval $(a, b) \subseteq U$ where $a < r < b$. These intervals must be disjoint, else they would not be maximal, and there are countably many of them since the rational numbers are countably infinite. On the real line $\mathbb{R}$ set $\mu(I) = b - a$ for all open intervals $I = (a, b)$. This extends uniquely by countable additivity (Meas5) to all open sets of $\mathbb{R}$. By finite additivity μ must take the value $b - a$ on all intervals, open, closed or half-open. For example, for any $\epsilon > 0$

$$(a-\epsilon, b) = (a-\epsilon, a) \cup [a, b),$$

and since the right-hand side is the union of two disjoint Borel sets we have

$$b - a + \epsilon = a - a + \epsilon + \mu\big([a, b)\big).$$

Hence the measure of the left-closed interval $[a, b)$ is

$$\mu\big([a, b)\big) = b - a.$$

Using

$$[a, b) = \{a\} \cup (a, b)$$

we see that every singleton has zero measure, $\mu(\{a\}) = 0$, and

$$(a,b] = (a,b) \cup \{b\},\ [a,b] = (a,b) \cup \{a\} \cup \{b\} \implies \mu\big((a,b]\big) = \mu\big([a,b]\big) = b - a.$$

Exercise: Show that if $A \subset \mathbb{R}$ is a countable set, then $\mu(A) = 0$.

Exercise: Show that the set of finite unions of left-closed intervals $[a, b)$ is closed with respect to the operation of taking differences of sets.

For any set $A \subset \mathbb{R}$ we define its **outer measure**

$$\mu^*(A) = \inf\{\mu(U) \mid U \text{ is an open set in } \mathbb{R} \text{ and } A \subseteq U\}. \tag{11.1}$$

While outer measure can be defined for arbitrary sets A of real numbers it is not really a measure at all, for it does not satisfy the countably additive property (Meas5). The best we can do is a property known as *countable subadditivity*: if $A_1, A_2, \ldots \in 2^X$ is any mutually disjoint sequence of sets then

$$\mu^*\Big(\bigcup_i A_i\Big) \le \sum_{i=1}^{\infty} \mu^*(A_i). \tag{11.2}$$

Proof: Let $\epsilon > 0$. For each $n = 1, 2, \ldots$ there exists an open set $U_n \supseteq A_n$ such that

$$\mu(U_n) \le \mu^*(A_n) + \frac{\epsilon}{2^n}.$$

Since $U = U_1 \cup U_2 \cup \ldots$ covers $A = A_1 \cup A_2 \cup \ldots$,

$$\mu^*(A) \le \mu(U) \le \mu(U_1) + \mu(U_2) + \cdots \le \sum_{i=1}^{\infty} \mu^*(A_n) + \epsilon.$$

As this is true for arbitrary $\epsilon > 0$, the inequality (11.2) follows immediately. ■

Exercise: Show that outer measure satisfies (Meas4), $\mu^*(\emptyset) = 0$.

Exercise: For an open interval $I = (a, b)$ show that $\mu^*(I) = \mu(I) = b - a$.

Exercise: Show that

$$A \subseteq B \implies \mu^*(A) \le \mu^*(B). \tag{11.3}$$

Following Carathéodory, a set E is said to be **Lebesgue measurable** if for any open interval $I = (a, b)$

$$\mu(I) = \mu^*(I \cap E) + \mu^*(I \cap E^c). \tag{11.4}$$

At first sight this may not seem a very intuitive notion.

What it is saying is that when we try to cover the mutually disjoint sets $I \cap E$ and $I - E = I \cap E^c$ with open intervals, the overlap of the two sets of intervals can be made ‘arbitrarily small’ (see Fig. 11.2). From now on we will often refer to Lebesgue measurable sets simply as *measurable*.

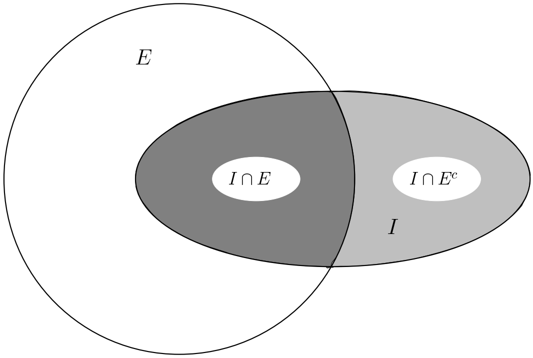

Figure 11.2 Lebesgue measurable set

Theorem 11.4 *If E is measurable then for any set A, measurable or not,*

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

Proof: Given $\epsilon > 0$, let U be an open set such that $A \subseteq U$ and

$$\mu(U) < \mu^*(A) + \epsilon. \tag{11.5}$$

Since $A = (A \cap E) \cup (A \cap E^c)$ is a union of two disjoint sets, Eq. (11.2) gives

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

Setting $U = \bigcup_n I_n$ where I_n are a finite or countable collection of disjoint open intervals, we have by (11.3)

$$\begin{aligned}\mu^*(A) &\le \mu^*(U \cap E) + \mu^*(U \cap E^c)\\ &\le \sum_n \mu^*(I_n \cap E) + \mu^*(I_n \cap E^c)\\ &= \sum_n \mu^*(I_n) \quad \text{since } E \text{ is measurable.}\end{aligned}$$

Using the inequality (11.5)

$$\mu^*(A) \le \mu^*(U \cap E) + \mu^*(U \cap E^c) < \mu^*(A) + \epsilon$$

for arbitrary $\epsilon > 0$. This proves the desired result. ■

Corollary 11.5 *If $E_1, E_2, \dots, E_n$ is any family of disjoint measurable sets then*

$$\mu^*(E_1 \cup E_2 \cup \dots \cup E_n) = \sum_{i=1}^{n} \mu^*(E_i).$$

Proof: If E and F are disjoint measurable sets, setting $A = E \cup F$ in (11.4) gives

$$\mu^*(E \cup F) = \mu^*\big((E \cup F) \cap E\big) + \mu^*\big((E \cup F) \cap E^c\big) = \mu^*(E) + \mu^*(F).$$

The result follows by induction on n. ■

Theorem 11.6 *The set of all Lebesgue measurable sets $\mathcal{L}$ is a σ-algebra, and μ^* is a measure on $\mathcal{L}$.*

Proof: The empty set is Lebesgue measurable, since for any open interval I,

$$\mu^*(I \cap \emptyset) + \mu^*(I \cap \mathbb{R}) = \mu^*(\emptyset) + \mu^*(I) = \mu^*(I) = \mu(I).$$

If E is a measurable set then so is E^c, on substituting in Eq. (11.4) and using $(E^c)^c = E$. Hence conditions (Meas1) and (Meas2) are satisfied. We prove (Meas3) in several stages.

Firstly, if E and F are measurable sets then $E \cup F$ is measurable. For, $I = \big(I \cap (E \cup F)\big) \cup \big(I \cap (E \cup F)^c\big)$ and using Eq. (11.2) gives

$$\mu^*(I) \le \mu^*\big(I \cap (E \cup F)\big) + \mu^*\big(I \cap (E \cup F)^c\big). \tag{11.6}$$

The sets in the two arguments on the right-hand side can be decomposed as

$$I \cap (E \cup F) = \big((I \cap F) \cap E\big) \cup \big((I \cap F) \cap E^c\big) \cup \big((I \cap F^c) \cap E\big)$$

and

$$I \cap (E \cup F)^c = I \cap (E^c \cap F^c) = (I \cap F^c) \cap E^c.$$

Again using Eq. (11.2) gives

$$\begin{aligned}&\mu^*\big(I \cap (E \cup F)\big) + \mu^*\big(I \cap (E \cup F)^c\big)\\ &\quad \le \mu^*\big((I \cap F) \cap E\big) + \mu^*\big((I \cap F) \cap E^c\big) + \mu^*\big((I \cap F^c) \cap E\big) + \mu^*\big((I \cap F^c) \cap E^c\big)\end{aligned}$$

On setting $A = I \cap F$ and $A = I \cap F^c$ respectively in Theorem 11.4 we have

$$\mu^*\big(I \cap (E \cup F)\big) + \mu^*\big(I \cap (E \cup F)^c\big) \le \mu^*(I \cap F) + \mu^*(I \cap F^c) = \mu^*(I),$$

since F is measurable. Combining this with the inequality (11.6) we conclude that for all intervals I

$$\mu^*\big(I \cap (E \cup F)\big) + \mu^*\big(I \cap (E \cup F)^c\big) = \mu^*(I) = \mu(I),$$

which shows that $E \cup F$ is measurable. Incidentally it also follows that the intersection $E \cap F = (E^c \cup F^c)^c$ and the difference $F - E = F \cap E^c$ is measurable. Simple induction shows that any finite union of measurable sets $E_1 \cup E_2 \cup \dots \cup E_n$ is measurable.

Let $E_1, E_2, \dots$ be any sequence of disjoint measurable sets. Set

$$S_n = \bigcup_{i=1}^{n} E_i \quad \text{and} \quad S = \bigcup_{i=1}^{\infty} E_i.$$

By subadditivity (11.2),

$$\mu^*(S) \le \sum_{i=1}^{\infty} \mu^*(E_i)$$

and since $S \supset S_n$ we have, using Corollary 11.5,

$$\mu^*(S) \geq \mu^*(S_n) = \sum_{i=1}^{n} \mu^*(E_i).$$

Since this holds for all integers n and the right-hand side is a monotone increasing series,

$$\mu^*(S) = \sum_{i=1}^{\infty} \mu^*(E_i). \tag{11.7}$$

If the series does not converge, the right-hand side is assigned the value ∞.

Since the E_i are disjoint sets, so are the sets $I \cap E_i$. Furthermore by Corollary 11.5

$$\sum_{i=1}^{n} \mu^*(I \cap E_i) = \mu^*\Big(\bigcup_{i=1}^{n}(I \cap E_i)\Big) = \mu^*\Big(I \cap \bigcup_{i=1}^{n} E_i\Big) \leq \mu^*(I).$$

Hence the series $\sum_{i=1}^{\infty} \mu^*(I \cap E_i)$ is convergent and for any $\epsilon > 0$ there exists an integer n such that

$$\sum_{i=n}^{\infty} \mu^*(I \cap E_i) < \epsilon.$$

Now since $S_n \subset S$,

$$I \cap S = I \cap \big(S_n \cup (S - S_n)\big) = (I \cap S_n) \cup \big(I \cap (S - S_n)\big)$$

and by subadditivity (11.2),

$$\begin{aligned}
\mu(I) &= \mu^*\big((I \cap S) \cup (I \cap S^c)\big) \\
&\leq \mu^*(I \cap S) + \mu^*(I \cap S^c) \\
&\leq \mu^*\big((I \cap S_n)\big) + \mu^*\big(I \cap (S - S_n)\big) + \mu^*\big(I \cap (S_n)^c\big) \\
&= \mu(I) + \mu^*\Big(I \cap \bigcup_{i=n+1}^{\infty} E_i\Big) \\
&= \mu(I) + \sum_{i=n+1}^{\infty} \mu^*(I \cap E_i) \\
&< \mu(I) + \epsilon.
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary

$$\mu(I) = \mu^*(I \cap S) + \mu^*(I \cap S^c),$$

which proves that S is measurable.

If $E_1, E_2, \ldots$ are a sequence of measurable sets, not necessarily disjoint, then set

$$F_1 = E_1,\ F_2 = E_2 - E_1, \ldots,\ F_n = E_n - (E_1 \cup E_2 \cup \cdots \cup E_{n-1}), \ldots$$

These sets are all measurable and disjoint, and

$$\bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} F_i.$$

The union of any countable collection $\{E_n\}$ of measurable sets is therefore measurable, proving that $\mathcal{L}$ is a σ-algebra. The outer measure μ^* is a measure on $\mathcal{L}$ since it satisfies $\mu^*(\emptyset) = 0$ and is countably additive by (11.7). ■

Theorem 11.6 shows that $(\mathbb{R}, \mathcal{L}, \mu = \mu^*\big|_{\mathcal{L}})$ is a measure space. The notation μ for Lebesgue measure agrees with the earlier convention $\mu\big((a, b)\big) = b - a$ on open intervals. All open sets are measurable as they are disjoint unions of open intervals. Since the Borel sets form the σ-algebra generated by all open sets, they are included in $\mathcal{L}$. Hence every Borel set is Lebesgue measurable. It is not true however that every Lebesgue measurable set is a Borel set.

A property that holds everywhere except on a set of measure zero is said to hold **almost everywhere**, often abbreviated to 'a.e.'. For example, two functions f and g are said to be equal a.e. if the set of points where $f(x) \neq g(x)$ is a set of measure zero. It is sufficient for the set to have outer measure zero, $\mu^*(A) = 0$, in order for it to have measure zero (see Problem 11.8).

Lebesgue measure is defined on cartesian product spaces $\mathbb{R}^n$ in a similar manner. We give the construction for $n = 2$. We have already seen that the σ-algebra of measurable sets on the product space $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ is defined as that generated by products of measurable sets $E \times E'$ where E and E' are Lebesgue measurable on $\mathbb{R}$. The outer measure of any set $A \subset \mathbb{R}^2$ is defined as

$$\mu^*(A) = \sup \sum_i \mu(I_i)\mu(I'_i)$$

where I_i and I'_i are any finite or countable family of open intervals of $\mathbb{R}$ such that the union $\bigcup_i I_i \times I'_i$ covers A. The outer measure of any product of open intervals is clearly the product of their measures, $\mu^*(I \times I') = \mu(I)\mu(I')$. We say a set $E \subset \mathbb{R}^2$ is **Lebesgue measurable** if, for any pair of open intervals I, I'

$$\mu(I)\mu(I') = \mu^*\big((I \times I') \cap E\big) + \mu^*\big((I \times I') \cap E^c\big).$$

As for the real line, outer measure μ^* is then a measure on $\mathbb{R}^2$. Lebesgue measure on higher dimensional products $\mathbb{R}^n$ is completely analogous. We sometimes denote this measure by μ^n.

Example 11.6 The Cantor set, Example 1.11, is a closed set since it is formed by taking the complement of a sequence of open intervals. It is therefore a Borel set and is Lebesgue measurable. The Cantor set is an uncountable set of measure 0 since the length remaining after the nth step in its construction is

$$1 - \tfrac{1}{3} - 2\big(\tfrac{1}{3}\big)^2 - 2^2\big(\tfrac{1}{3}\big)^3 - \cdots - 2^{n-1}\big(\tfrac{1}{3}\big)^n = \big(\tfrac{2}{3}\big)^n \to 0.$$

Its complement is an open subset of $[0, 1]$ with measure 1 – that is, having 'no gaps' between its component open intervals.

Example 11.7 Not every set is Lebegue measurable, but the sets that fail are non-constructive in character and invariably make use of the axiom of choice. A classic example is the following. For any pair of real numbers $x, y \in I = (0, 1)$ set $x\,Q\,y$ if and only if $x - y$

is a rational number. This is an equivalence relation on I and it partitions this set into disjoint equivalence classes $Q_x = \{y \mid y - x = r \in \mathbb{Q}\}$ where $\mathbb{Q}$ is the set of rational numbers. Assuming the axiom of choice, there exists a set T consisting of exactly one representative from each equivalence class Q_x. Suppose it has Lebesgue measure $\mu(T)$. For each rational number $r \in (-1, 1)$ let $T_r = \{x + r \mid x \in T\}$. Every real number $y \in I$ belongs to some T_r since it differs by a rational number r from some member of T. Hence, since $|r| < 1$ for each such T_r, we must have

$$(-1, 2) \supset \bigcup_r T_r \supset (0, 1).$$

The sets T_r are mutually disjoint and all have measure equal to $\mu(T)$. If the rational numbers are displayed as a sequence $r_1, r_2, \dots$ then

$$3 \geq \mu(T_1) + \mu(T_2) + \cdots = \sum_{i=1}^{\infty} \mu(T) \geq 1.$$

This yields a contradiction either for $\mu(T) = 0$ or $\mu(T) > 0$; in the first case the sum is 0, in the second it is ∞.

Problems

Problem 11.5 Show that every countable subset of $\mathbb{R}$ is measurable and has Lebesgue measure zero.

Problem 11.6 Show that the union of a sequence of sets of measure zero is a set of Lebesgue measure zero.

Problem 11.7 If $\mu^*(N) = 0$ show that for any set E, $\mu^*(E \cup N) = \mu^*(E - N) = \mu^*(E)$. Hence show that $E \cup N$ and $E - N$ are Lebesgue measurable if and only if E is measurable.

Problem 11.8 A measure is said to be **complete** if every subset of a set of measure zero is measurable. Show that if $A \subset \mathbb{R}$ is a set of outer measure zero, $\mu^*(A) = 0$, then A is Lebesgue measurable and has measure zero. Hence show that Lebesgue measure is complete.

Problem 11.9 Show that a subset E of $\mathbb{R}$ is measurable if for all $\epsilon > 0$ there exists an open set $U \supset E$ such that $\mu^*(U - E) < \epsilon$.

Problem 11.10 If E is bounded and there exists an interval $I \supset E$ such that

$$\mu^*(I) = \mu^*(I \cap E) + \mu^*(I - E)$$

then this holds for all intervals, possibly even those overlapping E.

Problem 11.11 The *inner measure* $\mu_*(E)$ of a set E is defined as the least upper bound of the measures of all measurable subsets of E. Show that $\mu_*(E) \leq \mu^*(E)$.

For any open set $U \supset E$, show that

$$\mu(U) = \mu_*(U \cap E) + \mu^*(U - E)$$

and that E is measurable with finite measure if and only if $\mu_*(E) = \mu^*(E) < \infty$.

11.3 Lebesgue integration

Let h be a simple function on a measure space $(X, \mathcal{M}, \mu)$,

$$h = \sum_{i=1}^{n} a_i \chi_{A_i} \quad (a_i \in \dot{\mathbb{R}})$$

where $A_i = h^{-1}(a_i)$ are measurable subsets of X. We define its **integral** to be

$$\int h \, \mathrm{d}\mu = \sum_{i=1}^{n} a_i \mu(A_i).$$

This integral only gives a finite answer if the measure of all sets A_i is finite, and in some cases it may not have a sensible value at all. For example $h : \mathbb{R} \to \mathbb{R}$ defined by

$$h(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

has integral $\infty + (-\infty)$ which is not well-defined.

If $h : X \to \mathbb{R}$ is a simple function, then for any constant b we have

$$\int bf \, \mathrm{d}\mu = b \int f \, \mathrm{d}\mu. \tag{11.8}$$

If $g = \sum_{j=1}^{m} b_j \chi_{B_j}$ is another simple function then $f + g$ is a simple function

$$f + g = \sum_{i=1}^{n} \sum_{j=1}^{m} (a_i + b_j) \chi_{A_i \cap B_j},$$

and has integral

$$\int (f + g) \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu + \int g \, \mathrm{d}\mu. \tag{11.9}$$

It is best to omit any term from the double sum where $a_i + b_j = 0$, else we may face the awkward problem of assigning a value to the product 0∞.

Exercise: Prove (11.8).

For any pair of functions $f, g : X \to \mathbb{R}$ we write $f \le g$ to mean $f(x) \le g(x)$ for all $x \in X$. If h and h' are simple functions such that $h \le h'$ then

$$\int h \, \mathrm{d}\mu \le \int h' \, \mathrm{d}\mu.$$

This follows immediately from the fact that $h - h'$ is a simple function that is non-negative everywhere, and therefore has an integral ≥ 0.

Taking the measure on $\mathbb{R}$ to be Lebesgue measure, the **integral of a non-negative measurable function** $f : X \to \mathbb{R}$ is defined as

$$\int f \, \mathrm{d}\mu = \sup \int h \, \mathrm{d}\mu$$

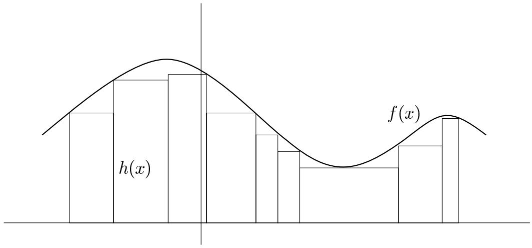

Figure 11.3 Integral of a non-negative measurable function

where the supremum is taken over all non-negative simple functions $h : X \to \mathbb{R}$ such that $h \le f$ (see Fig. 11.3). If $E \subset X$ is a measurable set then $f\chi_E$ is a measurable function on X that vanishes outside E. We define the **integral of f over E** to be

$$\int_E f \, d\mu = \int f\chi_E \, d\mu.$$

Exercise: Show that for any pair of non-negative measurable functions f and g

$$f \ge g \Longrightarrow \int f \, d\mu \ge \int g \, d\mu. \tag{11.10}$$

The following theorem is often known as the **monotone convergence theorem**:

Theorem 11.7 (Beppo Levi) *If f_n is an increasing sequence of non-negative measurable real-valued functions on X, $f_{n+1} \ge f_n$, such that $f_n(x) \to f(x)$ for all $x \in X$ then*

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

Proof: From the comments before Theorem 11.2 we know that f is a measurable function, as it is the limit of a sequence of measurable functions. If f has a finite integral then, by definition, for any $\epsilon > 0$ there exists a simple function $h : X \to \mathbb{R}$ such that $0 \le h \le f$ and $\int f \, d\mu - \int h \, d\mu < \epsilon$. For any real number $0 < c < 1$ let

$$E_n = \{x \in X \mid f_n(x) \ge ch(x)\} = (f_n - ch)^{-1}\big([0, \infty)\big),$$

clearly a measurable set for each positive integer n. Furthermore, since f_n is an increasing sequence of functions we have

$$E_n \subset E_{n+1} \quad \text{and} \quad X = \bigcup_{n=1}^{\infty} E_n,$$

since every point $x \in X$ lies in some E_n for n big enough. Hence

$$\int f \, d\mu \ge \int f_n \, d\mu \ge c \int h\chi_{E_n} \, d\mu.$$

If $h = \sum_i a_i \chi_{A_i}$ then

$$\int h\chi_{E_n} \, d\mu = \int \sum_i a_i \chi_{A_i \cap E_n} \, d\mu = \sum_i a_i \mu(A_i \cap E_n).$$

Hence, by Theorem 11.3,

$$\lim_{n\to\infty}\int h\chi_{E_n}\,\mathrm{d}\mu = \sum_i a_i\mu(A_i\cap X) = \sum_i a_i\mu(A_i) = \int h\,\mathrm{d}\mu,$$

so that

$$\int f\,\mathrm{d}\mu \ge \lim_{n\to\infty}\int f_n\,\mathrm{d}\mu \ge c\int h\,\mathrm{d}\mu \ge c\int f\,\mathrm{d}\mu - c\epsilon.$$

Since c can be chosen arbitrarily close to 1 and ϵ arbitrarily close to 0, we have

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

which proves the required result. ■

Exercise: How does this proof change if $\int f\,\mathrm{d}\mu = \infty$?

Using the result from Theorem 11.2, that every positive measurable function is a limit of increasing simple functions, it follows from Theorem 11.7 that simple functions can be replaced by arbitrary measurable functions in Eqs. (11.8) and (11.9).

Theorem 11.8 *The integral of a non-negative measurable function $f \ge 0$ vanishes if and only if $f(x) = 0$ almost everywhere.*

Proof: If $f(x) = 0$ a.e., let $h = \sum_i a_i\chi_{A_i} \ge 0$ be a simple function such that $h \le f$. Every set A_i must have measure zero, else $f(x) > 0$ on a set of positive measure. Hence $\int f\,\mathrm{d}\mu = \sup\int h\,\mathrm{d}\mu = 0$.

Conversely, suppose $\int f\,\mathrm{d}\mu = 0$. Let $E_n = \{x \mid f(x) \ge 1/n\}$. These are an increasing sequence of measurable sets, $E_{n+1} \supset E_n$, and

$$f \ge \frac{1}{n}\chi_{A_n}.$$

Hence

$$\int \frac{1}{n}\chi_{A_n}\,\mathrm{d}\mu = \frac{1}{n}\mu(A_n) \le \int f\,\mathrm{d}\mu = 0,$$

which is only possible if $\mu(A_n) = 0$. By Theorem 11.3 it follows that

$$\mu\big(\{x \mid f(x) > 0\}\big) = \mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \lim_{n\to\infty}\mu(A_n) = 0.$$

Hence $f(x) = 0$ almost everywhere. ■

Integration may be extended to real-valued functions that take on positive or negative values. We say a measurable function f is **integrable with respect to the measure** μ if both its positive and negative parts, f^+ and f^-, are integrable. The **integral of** f is then defined as

$$\int f\,\mathrm{d}\mu = \int f^+\,\mathrm{d}\mu - \int f^-\,\mathrm{d}\mu.$$

If f is integrable then so is its modulus $|f| = f^+ + f^-$, and

$$\left|\int f\,\mathrm{d}\mu\right| \le \left|\int f^+\,\mathrm{d}\mu - \int f^-\,\mathrm{d}\mu\right| \le \left|\int f^+\,\mathrm{d}\mu + \int f^-\,\mathrm{d}\mu\right| \le \int |f|\,\mathrm{d}\mu. \quad (11.11)$$

Hence a measurable function f is integrable if and only if $|f|$ is integrable.

A function $f : \mathbb{R} \to \mathbb{R}$ is said to be **Lebesgue integrable** if it is measurable and integrable with respect to the Lebesgue measure on $\mathbb{R}$. As for Riemann integration it is common to use the notations

$$\int f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x \equiv \int f\,\mathrm{d}\mu,$$

and for integration over an interval $I = [a, b]$,

$$\int_a^b f(x)\,\mathrm{d}x \equiv \int_I f\,\mathrm{d}\mu.$$

Riemann integrable functions on an interval $[a, b]$ are Lebesgue integrable on that interval. A function f is Riemann integrable if for any $\epsilon > 0$ there exist step functions h_1 and h_2 – simple functions that are constant on intervals – such that $h_1 \le f \le h_2$ and

$$\int_a^b h_2(x)\,\mathrm{d}x - \int_a^b h_1(x)\,\mathrm{d}x < \epsilon.$$

By taking $H_n = \sup(h_{i1})$ for the sequence of functions $(h_{i1}, h_{i2})\,(i = 1, \dots, n)$ defined by $\epsilon = 1, \frac{1}{2}, \dots, \frac{1}{n}$, it is straightforward to show that the H_n are simple functions the supremum of whose integrals is the Riemann integral of f. Hence f is Lebesgue integrable, and its Lebesgue integral is equal to its Riemann integral. The difference between the two concepts of integration is that for Lebesgue integration the simple functions used to approximate a function f need not be step functions, but can be constant on arbitrary measurable sets. For example, the function on [0, 1] defined by

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is irrational} \\ 0 & \text{if } x \text{ is rational} \end{cases}$$

is certainly Lebesgue integrable, and since $f = 1$ a.e. its Lebesgue integral is 1. It cannot, however, be approximated in the required way by step functions, and is not Riemann integrable.

Exercise: Prove the last statement.

Theorem 11.9 *If f and g are Lebesgue integrable real functions, then for any $a, b \in \mathbb{R}$ the function $af + bg$ is Lebesgue integrable and for any measurable set E*

$$\int_E (af + bg)\,\mathrm{d}\mu = a\int_E f\,\mathrm{d}\mu + b\int_E g\,\mathrm{d}\mu.$$

The proof is straightforward and is left as an exercise (see problems at end of chapter).

Lebesgue's dominated convergence theorem

One of the most important results of Lebesgue integration is that, under certain general circumstances, the limit of a sequence of integrable functions is integrable. First we need a lemma, relating to the concept of lim sup of a sequence of functions, defined in the paragraph prior to Theorem 11.2.

Lemma 11.10 (Fatou) *If (f_n) is any sequence of non-negative measurable functions defined on the measure space $(X, \mathcal{M}, \mu)$, then*

$$\int \liminf_{n\to\infty} f_n \, \mathrm{d}\mu \le \liminf_{n\to\infty} \int f_n \, \mathrm{d}\mu.$$

Proof: The functions $G_n(x) = \inf_{k\ge n} f_k(x)$ form an increasing sequence of non-negative measurable functions such that $G_n \le f_n$ for all n. Hence the lim inf of the sequence (f_n) is the limit of the sequence G_n,

$$\liminf_{n\to\infty} f_n = \sup_n G_n = \lim_{n\to\infty} G_n.$$

By the monotone convergence theorem 11.7,

$$\lim_{n\to\infty} \int G_n \, \mathrm{d}\mu = \int \liminf_{n\to\infty} f_n \, \mathrm{d}\mu$$

while the inequality $G_n \le f_n$ implies that

$$\int G_n \, \mathrm{d}\mu \le \int f_n \, \mathrm{d}\mu.$$

Hence

$$\int \liminf_{n\to\infty} f_n \, \mathrm{d}\mu = \lim_{n\to\infty} \int G_n \, \mathrm{d}\mu \le \liminf_{n\to\infty} \int f_n \, \mathrm{d}\mu.$$

■

Theorem 11.11 (Lebesgue) *Let (f_n) be any sequence of real-valued measurable functions defined on the measure space $(X, \mathcal{M}, \mu)$ that converges almost everywhere to a function f. If there exists a positive integrable function $g : X \to \mathbb{R}$ such that $|f_n| < g$ for all n then*

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

Proof: The function f is measurable since it is the limit a.e. of a sequence of measurable functions, and as $|f_n| < g$ all functions f_n and f are integrable with respect to the measure μ. Apply Fatou's lemma 11.10 to the sequence of positive measurable functions $g_n = 2g - |f_n - f| > 0$,

$$\int \liminf_{n\to\infty} (2g - |f_n - f|) \, \mathrm{d}\mu \le \liminf_{n\to\infty} \int (2g - |f_n - f|) \, \mathrm{d}\mu.$$

Since $\int g \, \mathrm{d}\mu < \infty$ and $\liminf |f_n - f| = \lim |f_n - f| = 0$, we have

$$0 \le \liminf_{n\to\infty} \int -|f_n - f| \, \mathrm{d}\mu = -\limsup_{n\to\infty} \int |f_n - f| \, \mathrm{d}\mu.$$

Since $|f_n - f| > 0$ this is only possible if

$$\lim_{n\to\infty} \int |f_n - f| \, d\mu = 0.$$

Hence

$$\left| \int (f_n - f) \, d\mu \right| \le \int |f_n - f| \, d\mu \to 0,$$

so that $\int f_n \, d\mu \to \int f \, d\mu$, as required. ■

The convergence in this theorem is said to be **dominated convergence**, g being the dominating function. An attractive feature of Lebesgue integration is that an integral over an unbounded set is defined exactly as for a bounded set. The same is true of unbounded integrands. This contrasts sharply with Riemann integration where such integrals are not defined directly, but must be defined as 'improper integrals' that are limits of bounded functions over a succession of bounded intervals. The concept of an improper integral is not needed at all in Lebesgue theory. However, Lebesgue's dominated convergence theorem can be used to evaluate such integrals as limits of finite integrands over finite regions.

Example 11.8 The importance of a dominating function is shown by the following example. The sequence of functions $(\chi_{[n,n+1]})$ consists of a 'unit hump' drifting steadily to the right and clearly has the limit $f(x) = 0$ everywhere. However it has no dominating function and the integrals do not converge

$$\int \chi_{[n,n+1]} \, d\mu = \int_n^{n+1} 1 \, dx = 1 \not\to \int 0 \, dx = 0.$$

We mention, without proof, the following theorem relating Lebesgue integration on higher dimensional Euclidean spaces to multiple integration.

Theorem 11.12 (Fubini) *If $f : \mathbb{R}^2 \to \mathbb{R}$ is a Lebesgue measurable function, then for each $x \in \mathbb{R}$ the function $f_x(y) = f(x, y)$ is measurable. Similarly for each $y \in \mathbb{R}$ the function $f'_y(x) = f(x, y)$ is measurable on $\mathbb{R}$. It is common to write*

$$\int f(x, y) \, dy \equiv \int f_x \, d\mu \quad \textit{and} \quad \int f(x, y) \, dx \equiv \int f'_y \, d\mu.$$

Then

$$\int\int f(x, y) \, dx \, dy \equiv \int f \, d\mu^2 = \int \Big(\int f(x, y) \, dy \Big) dx$$
$$= \int \Big(\int f(x, y) \, dx \Big) \, dy.$$

The result generalizes to a product of an arbitrary pair of measure spaces. For a proof see, for example, [1, 2].

Problems

Problem 11.12 Show that if f and g are Lebesgue integrable on $E \subset \mathbb{R}$ and $f \geq g$ a.e., then

$$\int_E f \, \mathrm{d}\mu \geq \int_E g \, \mathrm{d}\mu.$$

Problem 11.13 Prove Theorem 11.9.

Problem 11.14 If f is a Lebesgue integrable function on $E \subset \mathbb{R}$ then show that the function ψ defined by

$$\psi(a) = \mu(\{x \in E \mid |f(x)| > a\}) = O(a^{-1}) \qquad \textit{as } a \to \infty.$$

References

[1] N. Boccara. *Functional Analysis*. San Diego, Academic Press, 1990.

[2] B. D. Craven. *Lebesgue Measure and Integral*. Marshfield, Pitman Publishing Company, 1982.

[3] L. Debnath and P. Mikusiński. *Introduction to Hilbert Spaces with Applications*. San Diego, Academic Press, 1990.

[4] N. B. Haaser and J. A. Sullivan. *Real Analysis*. New York, Van Nostrand Reinhold Company, 1971.

[5] P. R. Halmos. *Measure Theory*. New York, Van Nostrand Reinhold Company, 1950.

[6] R. Geroch. *Mathematical Physics*. Chicago, The University of Chicago Press, 1985.

[7] C. de Witt-Morette, Y. Choquet-Bruhat and M. Dillard-Bleick. *Analysis, Manifolds and Physics*. Amsterdam, North-Holland, 1977.

[8] F. Riesz and B. Sz.-Nagy. *Functional Analysis*. New York, F. Ungar Publishing Company, 1955.