# A4 Integration (2017 HT)

Zhongmin Qian (Email: qianz@maths.ox.ac.uk)
Mathematical Institute, Oxford

December 3, 2016

## Contents

**1 Introduction** **3**

**2 Extended real line, upper and lower limits** **5**

**3 Measure spaces** **6**

**4 The Lebesgue measure** **8**
4.1 Outer measures and Carathéodory's extension theorem . . . 9
4.2 The Lebesgue measure space . . . 12
4.3 The Cantor set and null subsets . . . 14
4.4 An example of non measurable sets . . . 16
4.5 Lebesgue sub-spaces . . . 16

**5 Measurable functions** **17**
5.1 Definition and basic properties . . . 17
5.2 Almost everywhere properties . . . 21
5.3 Examples . . . 22

**6 Lebesgue integration** **24**
6.1 Lebesgue integrals, and integrable functions . . . 24
6.2 Integrals depending on parameters . . . 34
6.3 Examples . . . 35

**7 Fubini's theorem** **41**
7.1 Measures on product spaces . . . 41
7.2 Fubini's theorem and Tonelli's theorem . . . 43
7.3 Examples . . . 44
7.4 Change of variables . . . 47

**8 $L^p$-space and convergences** **49**
8.1 $L^p$-spaces . . . 50
8.2 Convergence in measure . . . 55
8.3 Convergence in $L^p$-space . . . 57

**9 Appendix 1 – From outer measures to measures** **58**

**10 Appendix 2 – Further results on Riemann integrals** **64**

These notes are written for the lectures delivered in 2016 Hilary term. In writing up these notes, I have benefited from the notes written by the previous lecturers, in particular the notes produced by A. Etheridge and C. Batty. I have adopted many of their worked examples, and in many places I have closely followed their notes. Any errors appeared in these notes however are completely my own responsibility. Please report errors, typos and etc. you may discover to `qianz@maths.ox.ac.uk` .

# 1 Introduction

In Prelims Analysis III – Integration, the theory of Riemann integrals for bounded continuous functions on finite intervals were established. In addition to the computational techniques such as the methods of change of variables, integration by parts and etc., we have proved several fundamental results about Riemann integrals. Let us recall two important results. The first is the existence of Riemann integrals. If $f$ is *bounded* on $[a,b]$ and is *continuous* on $(a,b)$ (where $a$ and $b$ are two numbers), then $f$ is Riemann integrable on $[a,b]$. The second is the fundamental theorem in calculus (FTC): if $F$ has *continuous derivative* on $[a,b]$ then

$$\int_a^b F'(x)dx = F(b) - F(a).$$

These results are important and are still good, and form the foundation for further study in your 3rd year and 4th year. What you have learned in Prelims Calculus and Analysis courses are important not only for the study of advanced theories in mathematics, but also are essential in applying mathematics to solve practical problems in science.

The theory of Riemann integration is a powerful tool for computations, it however lacks flexibility in handling limit procedures such as taking limits under integration. Here is a simple and illustrative example. Consider the Dirichlet function $f(x) = 1$ if $x$ is rational and $-1$ if $x$ is irrational. $f$ is a simple function but perhaps is not an interesting one. While it appears as the limit of simple functions. List all rational numbers in $[0,1]$ as $r_1$, $r_2$, $\cdots$, and define $f_n(x) = 1$ if $x \in \{r_1,\cdots,r_n\}$ and $f_n(x) = -1$ otherwise. Clearly $f_n \to f$ on $[0,1]$, clearly each $f_n$ is Riemann integrable on $[0,1]$ and $\int_0^1 f_n(x)dx = -1$. Therefore $\lim_{n\to\infty} \int_0^1 f_n(x)dx = -1$, but we can not take limit $\lim_n f_n = f$ first then take Riemann integral of $f$, as $f$ is not Riemann integrable.

It was recognized gradually that most difficulties one has with Riemann integrals come from the limitation of the theory of Riemann's integration. There was a need to extend the theory of integration to a larger class of functions, so that functions we are interested are integrable.

In order to explain the approach we are going to develop in this course, let us recall quickly the main steps in defining Riemann integrals. The first step is to choose a simple class of functions to which we can assign integrals. For the theory of Riemann integrals we choose the collection $\mathcal{L}$ of all *step functions*. Recall that a function $\varphi$ is step if $\varphi = \sum_{i=1}^n a_i 1_{J_i}$, where $n$ is a positive integer, $J_1$, $\cdots$, $J_n$ are *finite intervals*, and where $1_A$ denotes the characteristic function of $A$: $1_A(x) = 1$ or $0$ according to $x \in A$ or not. The integral of a step function $\varphi$ is defined to be $I(\varphi) = \sum_{i=1}^n a_i |J_i|$, where $|J|$ denotes the length of an interval $J$ (which is the *measure* of the interval $J$). The reason we choose to consider step functions in the Riemann integration lies in the fact that *the measure of an interval*, namely its length, makes sense for intervals. The second step is to define lower and upper integrals for a bounded function $f$ on a finite interval $(a,b)$. The lower integral $\underline{\int_a^b} f(x)dx$ is defined to be the supremum of all $I(\varphi)$ where $\varphi$ is step and $\varphi \leq f1_{(a,b)}$, and the upper integral $\overline{\int_a^b} f(x)dx$ is the infinimum of all $I(\psi)$ where $\psi$ is step and $\psi \geq f1_{(a,b)}$. Finally, $f$ is Riemann integrable if the lower and upper integrals of $f$ over $(a,b)$ coincide, and its Riemann integral $\int_a^b f(x)dx$ is the common upper (and lower) integral.

By definition of supremum and infinimum, if we are able to enlarge the class $\mathcal{L}$ of *simple* functions, then the corresponding lower integral $\underline{\int_a^b} f(x)dx$ (which is a supremum) would become greater, and the corresponding upper integral $\overline{\int_a^b} f(x)dx$ (an infinimum) become smaller, so they have better chance to be equal, thus have better chance to be "integrable". This suggests the following approach to extend the theory of integration to a larger class of functions. Take a collection of functions $\varphi$ in place of step functions with the following form

$$\varphi = \sum_{i=1}^{n} a_i 1_{E_i}$$

where $E_i$ are *certain subsets* but not necessary intervals. It is required that we should be able to define integral for such $\varphi$ to proceed the definition of integrals for general functions. Namely the integral of $\varphi$ should be defined as $I(\varphi) = \sum_{i=1}^{n} a_i |E_i|$. The problem to carry out this idea is that for a general subset $E$, $|E|$, the length of $E$, is not well-defined. The good news is that this is the main difficulty we need to overcome in order to establish a new integration theory. Therefore our first task is to extend the notion of lengths for intervals to a larger class of subsets than intervals, the new notion will be called the *Lebesgue measures*. As long as the measures of certain subsets are established, the integration theory can be established solely based on measures, no other structures of the real line will be needed, which is somehow an unexpected reward by extending the notion of length to measures to general sets beyond intervals.

*A very short history of Lebesgue's theory of integration.* E. Borel and H. Lebesgue were interested in the general construction of functions, and they wanted to extend the concepts of length, areas and volumes to general sets. They discovered it was not always possible to do so. It was H. Lebesgue who recognized the importance of the countable additivity in handling limits under integration. His Ph D thesis "Intégrale, Longueur, Aire" was finished in 1902, and in a book form "Lecons sur l'intégration et la recherche des fonctions primitives" which was published in 1904, in Paris, Lebesgue's theory of integration was established, and basic limit theorems (Monotone Convergence Theorem, Dominated Convergence Theorem), which are the powerful mathematical tools, were proved. H. Lebesgue applied his new integration theory to the study of trigonometric series, and published another monograph in 1906, "Lecons sur les séries trigonométriques" (Paris). The theory of Lebesgue's integration and its generalization, called *the theory of measures* (a measure is a generalization of the concept of length, area, volume), got prominent when, in 1933, A. Kolmogorov firmly established the foundation of probability theory by interpreting measure spaces as mathematical models for developing probability theory and statistics. A. Kolmogorov published his finding in the article "Grundbergriffe der Wahrscheinlichkeitsrechnung, *Erg. Mat.* 2, no.3 (1933). The English translation in a book form, "Foundations of the theory of probability" is still in print, and is still worthy of reading even today. A. Kolmogorov not only founded the probability theory based on the theory of measures, he made very important contributions even for the theory of measures. It was him who introduced the concept of conditional expectations which plays a vital role in both analysis and probability. He also developed basic tools for constructing measures out of marginal distributions.

Around 1940, J. L. Doob systematically developed the theory of martingales. He turned Kolmogorov's conditional expectation into a powerful tool, and identified a class of random variables (while, random variables are exactly those measurable functions according to A. Kolmogorov) called martingales which is the most important class of measurable functions in probability theory. Doob established powerful tools in a collection of Doob's inequalities, convergence theorems and etc. His results on martingales were organized neatly in his classic "Stochastic Processes", published in 1953. Meanwhile Paul Halmos wrote a book on the subject of measures, "Measure Theory" (1950), which was and remains a standard reference on Lebesgue's theory of measures and integration.

# 2 Extended real line, upper and lower limits

We will use the following convention in dealing with the symbols $\infty$ and $-\infty$. Introduce the extended real line $[-\infty, \infty] = \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$ by adding two symbols $-\infty$ and $\infty$ which are not in $\mathbb{R}$. We make the following conventions: for every $a \in \mathbb{R}$, $-\infty < a < \infty$,

$$a + \infty = \infty + a = \infty, \;\; a - \infty = -\infty + a = -\infty.$$

$$\frac{a}{-\infty} = \frac{a}{\infty} = 0, \text{ and } 0 \cdot \infty = -\infty \cdot 0 = 0,$$

but $\frac{\infty}{\infty}$ and $\frac{a}{0}$ are not defined.

Recall that if $S \subseteq \mathbb{R}$ is non-empty, and if $S$ is bounded above, that is, there is $b \in \mathbb{R}$ such that $s \leq b$ for every $s \in S$, then the Completeness Axiom says that $S$ has the least upper bound denoted by $\sup S$, which is an upper bound of $S$, and if $b$ is an upper bound of $S$, then $\sup S \leq b$. If $S$ is not bounded above, that is, for every $n = 1, 2, \cdots$, there is $s_n \in S$, such that $s_n > n$, then we say $\sup S = \infty$. Clearly, if $S \subseteq \mathbb{R}$ is non-empty, then $\sup S = \infty$ if and only if there is a sequence $\{s_n\}$, where each $s_n \in S$, such that $s_n \to \infty$. Similarly we extend the definition of $\inf S$ for any non-empty subset $S$ of $\mathbb{R}$. It is remains true that $\inf S = -\sup(-S)$, where $-(\infty) = -\infty$ which is a convention, and $-S = \{-x : x \in S\}$, for non-empty $S \subseteq \mathbb{R}$.

If $\{a_n\}$ is a sequence of real numbers, then, by our extension of the concept for sup and inf,

$$\sup_{k \geq n} a_k = \lim_{m \to \infty} \max \{a_n, a_{n+1}, \cdots, a_{n+m}\},$$

which is decreasing in $n$, and

$$\inf_{k \geq n} a_k = \lim_{m \to \infty} \min \{a_n, a_{n+1}, \cdots, a_{n+m}\}$$

which is increasing in $n$. Define

$$\liminf_{n \to \infty} a_n = \lim_{n \to \infty} \lim_{m \to \infty} \min \{a_{n+1}, \cdots, a_{n+m}\}$$

and

$$\limsup_{n \to \infty} a_n = \lim_{n \to \infty} \lim_{m \to \infty} \max \{a_{n+1}, \cdots, a_{n+m}\},$$

which are called the lower and upper limit of $\{a_n\}$, respectively.

It follows that $\liminf_{n \to \infty} a_n = -\limsup_{n \to \infty}(-a_n)$ and $\limsup_{n \to \infty} a_n = -\liminf_{n \to \infty}(-a_n)$.

It also follows from definition that, which are useful to evaluate upper and lower limits, $\liminf_{n \to \infty} a_n$ is the least number among all possible limits (i.e. numbers or $\pm\infty$ which are the limits of convergent sub-sequences of $\{a_n\}$), and similarly, $\limsup_{n\to\infty} a_n$ is the largest number among all possible limits (i.e. numbers or $\pm\infty$ which are the limits of convergent sub-sequences of $\{a_n\}$).

If $\{b_n\}$ is a sequence of real numbers, and if $\lim_{n\to\infty} b_n = b$ exists and finite (i.e. $b \neq \infty$ or $-\infty$), then

$$\liminf_{n\to\infty}(a_n + b_n) = \liminf_{n\to\infty} a_n + \lim_{n\to\infty} b_n$$

and

$$\limsup_{n\to\infty}(a_n + b_n) = \limsup_{n\to\infty} a_n + \lim_{n\to\infty} b_n.$$

# 3 Measure spaces

Our first task is to extend the notion of length to some subsets $E$ of $\mathbb{R}$, called Lebesgue measures and denoted by $m(E)$ (which is the length $|J|$ if $E = J$ is an interval). Suppose the class of subsets of $\mathbb{R}$ to which we are able to assign measures is denoted by $\mathcal{M}$, then we expect the followings are satisfied:

1. The empty set $\emptyset$ has measure zero, that is $m(\emptyset) = 0$.

2. The measure of $E \in \mathcal{M}$ is non-negative, i.e. $m(E) \geq 0$ (but can be infinity).

3. *$m$ is finite additive*: if $E_1, \cdots, E_n$ belong to $\mathcal{M}$, then $E_1 \cup \cdots \cup E_n \in \mathcal{M}$, and if in addition $E_i$ are disjoint, then

$$m(E_1 \cup \cdots \cup E_n) = m(E_1) + \cdots + m(E_n).$$

   The property of finite additivity is not enough in order to do integration, which should be enhanced.

4. *$m$ is countably additive*: if $E_1, \cdots, E_n, \cdots$ belong to $\mathcal{M}$, then $\cup_{n=1}^{\infty} E_n \in \mathcal{M}$. If in addition $\{E_n : n = 1, 2, \cdots\}$ are disjoint, then

$$m\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} m(E_n).$$

We are going to identify carefully the class $\mathcal{M}$ of *Lebesgue measurable subsets* of $\mathbb{R}$, and to define the measure $m$ as a function from $\mathcal{M}$ to $[0, \infty]$ which possesses properties 1)-4).

The theory of integration based on the Lebesgue measure $m$ may be developed in a rather general setting which uses no algebraic or geometric structures of $\mathbb{R}$. Therefore it is beneficial to introduce the concept of *measures*, the concept of *measurable spaces*, and the concept of *measure spaces.*

**Definition 3.1** *Let $\Omega$ be a set [which is called a* space *or called a* sample space*], and $\mathcal{F}$ be a collection of some subsets of $\Omega$.*

*1) $\mathcal{F}$ is called an* algebra *on $\Omega$ if a) $\emptyset \in \mathcal{F}$ and $\Omega \in \mathcal{F}$, b) if $A, B \in \mathcal{F}$ then $A^c \in \mathcal{F}$ and $A \cup B \in \mathcal{F}$.*

*2) $\mathcal{F}$ is called a* $\sigma$-algebra *(also called a* $\sigma$-field*) on $\Omega$, if $\mathcal{F}$ is an algebra over $\Omega$, and if in addition $\mathcal{F}$ is closed under the countable union operation. That is, if $A_1, A_2, \cdots$ belong to $\mathcal{F}$, so does $\cup_{n=1}^{\infty} A_n$.*

*3) A pair $(\Omega, \mathcal{F})$, where $\mathcal{F}$ is a $\sigma$-algebra on $\Omega$, is called a* measurable space. *If $E \in \mathcal{F}$ then $E$ is called a measurable subset of $\Omega$ (with respect to the $\sigma$-algebra $\mathcal{F}$).*

**Remark 3.2** *(about notations) 1) If $A$ is a subset of $\Omega$ then $A^c$ denotes $\Omega \setminus A$ if no confusion may arise, so $A^c = \{x \in \Omega : x \notin A\}$.*

*2) Some authors use $AB$ to denote $A \cap B$, in particular in probability literature. The notation is justified since $1_{A \cap B} = 1_A 1_B$, a fact which will be used without further comments.*

*3) According to De Morgan's law, if $\mathcal{F}$ is a $\sigma$-algebra and $A_n \in \mathcal{F}$ $(n = 1, 2, \cdots)$ then $\bigcap_{n=1}^{\infty} A_n = (\bigcup_{n=1}^{\infty} A_n^c)^c \in \mathcal{F}$.*

**Definition 3.3** *Let $(\Omega, \mathcal{F})$ be a measurable space. A function $\mu : \mathcal{F} \to [0, \infty]$ is called a* measure *on $(\Omega, \mathcal{F})$, if*

*1) $\mu(\emptyset) = 0$ and $\mu(A) \in [0, \infty]$ (i.e. $\mu(A) \geq 0$ or $\mu(A) = \infty$) for every $A \in \mathcal{F}$.*

*2) [$\mu$ is countably additive] If $A_n \in \mathcal{F}$ for $n = 1, 2, \cdots$, and $A_n$ are disjoint, then $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$.*

*A triple $(\Omega, \mathcal{F}, \mu)$, where $\mathcal{F}$ is a $\sigma$-algebra $\Omega$ and $\mu$ is a measure on the measurable space $(\Omega, \mathcal{F})$, is called a* measure space.

A measure $\mu$ on a measurable space $(\Omega, \mathcal{F})$ is called a *probability* (measure) if $\mu(\Omega) = 1$. In this case $\Omega$ is called a sample space (of fundamental events), an element $A$ in the $\sigma$-algebra $\mathcal{F}$ is called an *event*, and $\mu(A)$ is called the probability that the event $A$ happens.

**Proposition 3.4** *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Then*

*1) $\mu(A) \leq \mu(B)$ if $A \subset B$,*

*2) if $A_n \in \mathcal{F}$ and $A_n \uparrow$ (that is $A_n \subset A_{n+1}$ for all $n$), then $\mu(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mu(A_n)$,*

*3) if $A_n \in \mathcal{F}$, $A_n \downarrow$ (that is $A_n \supset A_{n+1}$ for all $n$) and $\mu(A_1) < \infty$, then $\mu(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mu(A_n)$.*

**Proof.** 1) If $A \subset B$ then $B = A \cup (B \cap A^c)$, and $A$ and $B \cap A^c$ are disjoint, by additivity of the measure, we have

$$\mu(B) = \mu(A) + \mu(B \cap A^c) \geq \mu(A).$$

2) Let $E_1 = A_1$ and $E_n = A_n - A_{n-1}$ for $n \geq 2$. Then $E_n$ are disjoint and $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} E_n$. Therefore, by the countable additivity of $\mu$,

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

On the other hand, since $A_n \uparrow$, $\bigcup_{k=1}^{n} E_k = A_n$, thus

$$\sum_{k=1}^{\infty} \mu(E_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(E_k) = \lim_{n\to\infty} \mu\left[\bigcup_{k=1}^{n} E_k\right] = \lim_{n\to\infty} \mu(A_n)$$

which yields 2).

3) Let $B_n = A_1 \setminus A_n$. Then $B_n \uparrow$. Since $A_1 = B_n \cup A_n$, $B_n$ and $A_n$ are disjoint, so $\mu(A_1) = \mu(B_n) + \mu(A_n)$. Hence $\mu(B_n) = \mu(A_1) - \mu(A_n)$ as $\mu(A_n) \leq \mu(A_1) < \infty$. According to de Morgan's law $\bigcup_{n=1}^{\infty} B_n = A_1 \setminus \bigcap_{n=1}^{\infty} A_n$ , thus, by applying 2) to $B_n$, we obtain that

$$\begin{aligned} \mu(A_1) - \mu\left[\bigcap_{n=1}^{\infty} A_n\right] &= \mu\left[A_1 \setminus \bigcap_{n=1}^{\infty} A_n\right] = \mu\left[\bigcup_{n=1}^{\infty} B_n\right] \\ &= \lim_{n\to\infty} \mu\left(B_n\right) = \lim_{n\to\infty} \left(\mu(A_1) - \mu(A_n)\right) \\ &= \mu(A_1) - \lim_{n\to\infty} \mu(A_n) \end{aligned}$$

which implies 3) for $\mu(A_1) < \infty$. ■

Interesting examples of measures are Lebesgue's measures on Euclidean spaces $\mathbb{R}^d$ which will be constructed in the next section. One of the central problems in probability theory, quantum field theories and statistical mechanics is to construct various measures on some infinite dimensional spaces, which however should be studied in specialized courses.

# 4 The Lebesgue measure

In this section we construct 1) the $\sigma$-algebra $\mathcal{M}_{\text{Leb}}$ of Lebesgue measurable subsets of $\mathbb{R}$, and 2) the Lebesgue measure $m : \mathcal{M}_{\text{Leb}} \to [0, \infty]$.

Let us first describe a candidate of the Lebesgue measure $m$, called the Lebesgue outer measure.

Let $\mathcal{C}$ denote the collection of all finite intervals with a form $(a, b]$ where $a \leq b$ are two real numbers. $\mathcal{C}$ is a $\pi$-system over $\mathbb{R}$ in the sense that if $A, B \in \mathcal{C}$ then $A \cap B \in \mathcal{C}$. If $J$ is an interval, then $|J|$ denotes the length of $J$, so if $J = (a, b] \in \mathcal{C}$ then $|J| = b - a$.

We build the *outer measure* $m^*$ by

$$m^*(A) = \inf\left\{\sum_{i=1}^{\infty} |J_i| : \text{ where all } J_i \in \mathcal{C} \text{ such that } \bigcup_{i=1}^{\infty} J_i \supseteq A\right\} \tag{4.1}$$

for $A \subset \mathbb{R}$. Then $m^*$ possesses the following properties:

(1) $m^*(\emptyset) = 0$, and $m^*(A) \geq 0$ for any $A \subseteq \mathbb{R}$.

(2) $m^*(A) \leq m^*(B)$ if $A \subset B$, and

(3) $m^*$ is countably *sub-additive*, as stated in the following

**Lemma 4.1** [Countably additive] *If $\{A_n : n = 1, 2, \cdots\}$ is a sequence of subsets, then*

$$m^*\left[\bigcup_{n=1}^{\infty} A_n\right] \leq \sum_{n=1}^{\infty} m^*(A_n). \tag{4.2}$$

**Proof.** If $\sum_{n=1}^{\infty} m^*(A_n) = \infty$ then (4.2) holds. Suppose $\sum_{n=1}^{\infty} m^*(A_n) < \infty$. Then $m^*(A_n) < \infty$ for every $n$. By definition of $m^*(A_n)$, for every $\varepsilon > 0$ there is a countable cover $\left\{J_i^{(n)} : i = 1, 2, \cdots\right\}$ of $A_n$, where $J_i^{(n)} \in \mathcal{C}$, such that

$$\sum_{i=1}^{\infty} |J_i^{(n)}| \leq m^*(A_n) + \frac{\varepsilon}{2^n}.$$

While $\left\{J_i^{(n)} : i, n = 1, 2, \cdots\right\}$ forms a countable cover of $\bigcup_{n=1}^{\infty} A_n$ so that

$$\begin{aligned} m^* \bigcup_{n=1}^{\infty} A_n &\leq \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} |J_i^{(n)}| \leq \sum_{n=1}^{\infty} \left( m^*(A_n) + \frac{\varepsilon}{2^n} \right) \\ &= \sum_{n=1}^{\infty} m^*(A_n) + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, (4.2) follows immediately. ■

Although $m^*(E)$ is well defined for *every subset* $E$ of $\mathbb{R}$, $m^*$ is not countably additive, thus $m^*$ is not a measure on $\mathcal{P}(\mathbb{R})$ (the $\sigma$-algebra of all subsets of $\mathbb{R}$). In fact, $\mathcal{P}(\mathbb{R})$ is too big for $m^*$ to be countably additive. The main technical step in the construction of the Lebesgue measure is to identify the $\sigma$-algebra $\mathcal{M}_{\text{Leb}}$ on which $m^*$ is countably additive. $\mathcal{M}_{\text{Leb}}$ should be sufficient large and should include all intervals. This will be achieved in the celebrated Carathéodory's extension theorem.

## 4.1 Outer measures and Carathéodory's extension theorem

This is a major theorem in the theory of measures. It is a general theorem so we will formulate it in general setting. *Its proof is not examinable.* We will use this theorem to identify some important measurable subsets which we should be familiar with.

Let $(\Omega, \mathcal{G})$ be a measurable space, and $\mu^* : \mathcal{G} \to [0, \infty]$ be an *outer measure* on $(\Omega, \mathcal{G})$ in the following sense:

1) $\mu^*(\varnothing) = 0$, and $\mu^*(A) \geq 0$ for every $A \in \mathcal{G}$;
2) $\mu^*(A) \leq \mu^*(B)$ if $A \subseteq B$, $A, B \in \mathcal{G}$;
3) $\mu^*$ is *countably sub-additive*: if $A_n \in \mathcal{G}$ for $n = 1, 2, \cdots$, then

$$\mu^* \left[ \bigcup_{n=1}^{\infty} A_n \right] \leq \sum_{n=1}^{\infty} \mu^*(A_n). \tag{4.3}$$

**Definition 4.2** *A subset $E \in \mathcal{G}$ is $\mu^*$-measurable if $E$ satisfies the* Carathéodory condition

$$\mu^*(F) = \mu^*(F \cap E) + \mu^*(F \cap E^c) \quad \textit{for every } F \in \mathcal{G}. \tag{4.4}$$

*The collection of all $\mu^*$-measurable subsets is denoted by $\mathcal{G}^m$.*

Since $F \setminus E = F \cap E^c$, (4.4) may be written as

$$\mu^*(F) = \mu^*(F \cap E) + \mu^*(F \setminus E) \quad \text{for any subset } F \in \mathcal{G}. \tag{4.5}$$

Since $F = (F \cap E) \cup (F \cap E^c)$, by sub-additivity of $\mu^*$

$$\mu^*(F) \leq \mu^*(F \cap E) + \mu^*(F \cap E^c)$$

for any $E$ and $F$ belonging to $\mathcal{G}$ , the Carathéodory condition (4.4) is equivalent to the inequality

$$\mu^*(F) \geq \mu^*(F \cap E) + \mu^*(F \cap E^c) \tag{4.6}$$

for any subset $F \in \mathcal{G}$.

$E$ is measurable if and only if (4.6) holds for any subset $F \in \mathcal{G}$ such that $\mu^*(F) < \infty$.

**Theorem 4.3** *$\mathcal{G}^m$ is a $\sigma$-algebra on $\Omega$, and $\mu^*$ is a measure on $\mathcal{G}^m$.*

**Proof.** *[The proof is not examinable].* Clearly the empty set $\varnothing \in \mathcal{G}^m$. Also, according to (4.5), $E \in \mathcal{G}^m$ if and only if $E^c \in \mathcal{G}^m$.

Let $A, B \in \mathcal{G}^m$. We show that $A \cap B \in \mathcal{G}^m$ so that $\mathcal{G}^m$ is an algebra. Since $(A \cap B)^c = A^c \cup B^c$, we need to show that

$$\mu^*(F) = \mu^*(F \cap (A \cap B)) + \mu^*(F \cap (A^c \cup B^c))$$

for any subset $F \in \mathcal{G}$.

Since $A$ is $\mu^*$-measurable, so

$$\mu^*(F) = \mu^*(F \cap A) + \mu^*(F \cap A^c). \tag{4.7}$$

Since $B$ is also $\mu^*$-measurable, applying (4.5) to $F \cap A$ (in the place of $F$) and $B$ we obtain

$$\mu^*(F \cap A) = \mu^*(F \cap A \cap B) + \mu^*(F \cap A \cap B^c). \tag{4.8}$$

Substitute (4.8) into (4.7) to obtain

$$\mu^*(F) = \mu^*(F \cap (A \cap B)) + \mu^*(F \cap A \cap B^c) + \mu^*(F \cap A^c). \tag{4.9}$$

Use again (4.5) to $A$ (which is measurable) and $F \cap (A^c \cup B^c)$ to obtain

$$\begin{aligned} \mu^*(F \cap (A^c \cup B^c)) &= \mu^*(F \cap (A^c \cup B^c) \cap A) + \mu^*(F \cap (A^c \cup B^c) \cap A^c) \\ &= \mu^*(F \cap B^c \cap A) + \mu^*(F \cap A^c), \end{aligned}$$

here we have used the elementary equalities $(A^c \cup B^c) \cap A = B^c \cap A$ and $(A^c \cup B^c) \cap A^c = A^c$. Together with (4.9) we deduce that

$$\mu^*(F) = \mu^*(F \cap (A \cap B)) + \mu^*(F \cap (A^c \cup B^c))$$

for any $F \in \mathcal{G}$, so that $A \cap B \in \mathcal{G}^m$. By de Morgan law, it follows also that

$$A \cup B = (A^c \cap B^c)^c \in \mathcal{G}^m.$$

Thus $\mathcal{G}^m$ is an algebra.

We next to show that $\mathcal{G}^m$ is a $\sigma$-algebra. To this end, consider $E_n \in \mathcal{G}^m$, $n = 1, 2, \cdots$. We show that $E = \bigcup_{n=1}^\infty E_n \in \mathcal{G}^m$. Without losing generality we may assume that $E_n$ are disjoint, otherwise we may consider $A_n$ instead, where $A_1 = E_1 \in \mathcal{G}^m$ and $A_n = E_n \setminus (\cup_{j<n} E_j) \in \mathcal{G}^m$ for $n \geq 2$. $A_n \in \mathcal{G}^m$ are disjoint, and $E = \bigcup_{n=1}^\infty A_n$.

Since $\bigcup_{j=1}^n E_j \in \mathcal{G}^m$, by applying (4.5), we obtain

$$\mu^*(F) = \mu^*\left[F \cap \left(\bigcup_{j=1}^n E_j\right)\right] + \mu^*\left[F \cap \left(\bigcup_{j=1}^n E_j\right)^c\right].$$

Since $E_n$ is measurable,

$$\begin{aligned}
\mu^*\left[F\cap\left(\bigcup_{j=1}^n E_j\right)\right] &= \mu^*\left[F\cap\left(\bigcup_{j=1}^n E_j\right)\cap E_n\right]+\mu^*\left[F\cap\left(\bigcup_{j=1}^n E_j\right)\cap E_n^c\right] \\
&= \mu^*\left[F\cap E_n\right]+\mu^*\left[F\cap\left(\bigcup_{j=1}^{n-1} E_j\right)\right] \\
&= \cdots = \sum_{j=1}^n \mu^*\left[F\cap E_j\right]
\end{aligned}$$

here we have used the following identities: since $E_j$ are disjoint

$$\left(\bigcup_{j=1}^n E_j\right)\cap E_n = E_n \text{ and } \left(\bigcup_{j=1}^n E_j\right)\cap E_n^c = \bigcup_{j=1}^{n-1} E_j.$$

We therefore have

$$\begin{aligned}
\mu^*(F) &= \sum_{j=1}^n \mu^*\left[F\cap E_j\right]+\mu^*\left[F\cap\left(\bigcup_{j=1}^n E_j\right)^c\right] \\
&\geq \sum_{j=1}^n \mu^*(F\cap E_j)+\mu^*\left[F\cap\left(\bigcup_{j=1}^\infty E_j\right)^c\right].
\end{aligned}$$

Letting $n\to\infty$ to obtain

$$\begin{aligned}
\mu^*(F) &\geq \sum_{j=1}^\infty \mu^*(F\cap E_j)+\mu^*\left[F\cap\left(\bigcup_{j=1}^\infty E_j\right)^c\right] && (4.10) \\
&\geq \mu^*\left[F\cap\left(\bigcup_{j=1}^\infty E_j\right)\right]+\mu^*\left[F\cap\left(\bigcup_{j=1}^\infty E_j\right)^c\right] && (4.11)
\end{aligned}$$

which implies that $\bigcup_{j=1}^\infty E_j \in \mathcal{G}^m$. Since we must have

$$\mu^*(F)\leq \mu^*\left[F\cap\left(\bigcup_{j=1}^\infty E_j\right)\right]+\mu^*\left[F\cap\left(\bigcup_{j=1}^\infty E_j\right)^c\right]$$

for any $F\in\mathcal{G}$ and $E_j$, the inequalities in (4.11) must be equalities. Hence

$$\mu^*(F)=\sum_{j=1}^\infty \mu^*\left[F\cap E_j\right]+\mu^*\left[F\cap\left(\bigcup_{j=1}^\infty E_j\right)^c\right]$$

for any subset $F\in\mathcal{G}$. In particular, by applying the equality above to $F=\bigcup_{k=1}^\infty E_k$ we obtain

$$\mu^*\left[\bigcup_{j=1}^\infty E_j\right]=\sum_{j=1}^\infty \mu^*\left[\left(\bigcup_{k=1}^\infty E_k\right)\cap E_j\right]=\sum_{j=1}^\infty \mu^*(E_j).$$

That is, $\mu^*$ is countably additive, so that $\mu^*$ is a measure on $\mathcal{G}^m$. ∎

It is possible, for some outer measures $\mu^*$, the $\sigma$-algebra $\mathcal{G}^m$ may be trivial, so that it is useless in this case.

Let us look at some simple examples of $\mu^*$-measurable sets.

**Definition 4.4** *A subset $A \in \mathcal{G}$ is called a $\mu^*$-*null set *if it has zero $\mu^*$-outer measure, i.e. $\mu^*(A) = 0$.*

We have the following simple fact

**Lemma 4.5** *If $\{A_n : n = 1, 2, \cdots\}$ is a sequence of $\mu^*$-null subsets, then $\bigcup_{n=1}^{\infty} A_n$ is $\mu^*$-null as well.*

This follows immediately from the fact that $\mu^*$ is countably sub-additive. In fact, if $\mu^*(A_n) = 0$, then

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n) = 0$$

hence we must have $\mu^*\left[\bigcup_{n=1}^{\infty} A_n\right] = 0$. $\bigcup_{n=1}^{\infty} A_n$ is $\mu^*$-null.

**Proposition 4.6** *If $A$ is $\mu^*$-null, then $A \in \mathcal{G}^m$.*

**Proof.** For every subset $F \in \mathcal{G}$, we have $\mu^*(F \cap A) \leq \mu^*(A) = 0$ so that

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

By definition, $A$ is $\mu^*$-measurable. ∎

In applications, it is important to be able to describe the structure of the $\sigma$-algebra $\mathcal{G}^m$ of $\mu^*$-measurable subsets, which is determined by the outer measure $\mu^*$. However, without further information about $\mu^*$, Prop. 4.6 is the best we can offer. In practice, we do have some a priori knowledge, which helps to identify the structure of $\mathcal{G}^m$.

## 4.2 The Lebesgue measure space

Recall that $\mathscr{C}$ is the $\pi$-system of all intervals $(a, b]$ where $a \leq b$ are two real numbers. If $J = (a, b]$, then $|J| = b - a$ is the length of $J$. The Lebesgue outer measure $m^*$ is defined by

$$m^*(E) = \inf\left\{\sum_{i=1}^{\infty} |J_i| : J_i \in \mathscr{C} \text{ such that } \bigcup_{i=1}^{\infty} J_i \supset E\right\} \tag{4.12}$$

where $E$ is a subset of $\mathbb{R}$, and the inf takes over *all possible countable* cover $\{J_i : i = 1, 2, \cdots\} \subset \mathscr{C}$ of $E$.

$\mathcal{M}_{\text{Leb}}$ denotes the collection of all $m^*$-measurable [from now on, called Lebesgue measurable, or simply measurable if no confusion may arise] subsets. Recall that $E$ is $m^*$-measurable if $E$ satifies the Caratheodory condition:

$$m^*(F) = m^*(F \cap E) + m^*(F \cap E^c)$$

for every $F \subseteq \mathbb{R}$.

We have proved that $\mathcal{M}_{\text{Leb}}$ is a $\sigma$-algebra, and $m^*$ restricted on $\mathcal{M}_{\text{Leb}}$ is a measure. The measure space $(\mathbb{R}, \mathcal{M}_{\text{Leb}}, m^*)$ is called the Lebesgue measure space, or simply the Lebesgue space.

We give a description of the sets in $\mathcal{M}_{\text{Leb}}$ in this lecture. The key is the following

**Lemma 4.7** *$\mathscr{C} \subseteq \mathcal{M}_{Leb}$, and moreover, if $J = (a, b] \in \mathscr{C}$, then $m^*(J) = b - a$.*

This lemma looks very simple, but its proof is not easy and is quite technical. Its proof is not examinable, therefore I don't give the proof in the lecture.

By using this lemma, we can describe the $\sigma$-algebra $\mathcal{M}_{\text{Leb}}$. We divide this task into several steps, all are elementary.

1) *Firstly, we claim any interval $J$ is Lebesgue measurable, and $m^*(J) = |J|$ the length of $J$.*

If $J$ is unbounded, for example $J = (a, \infty)$, then $J = \bigcup_{n=1}^{\infty}(a, n]$, since each $(a, n]$ is measurable and has measure $n - a$, so that $J \in \mathcal{M}_{\text{Leb}}$, and $m^*(J) \geq n - a$, therefore $m^*(J) = \infty$. If $J = [a, \infty)$, then $J = \{a\} \cup (a, \infty)$. We only need to show $\{a\} = [a, a]$ is measurable. Indeed $[a, a] = \bigcap_{n=1}^{\infty}(a - \frac{1}{n}, a]$, so $[a, a] \in \mathcal{M}_{\text{Leb}}$ and

$$0 \leq m^*(\{a\}) \leq a - (a - \frac{1}{n}) = \frac{1}{n}$$

for any $n = 1, 2, \cdots$, so $m^*(\{a\}) = 0$.

While $(-\infty, a) = [a, \infty)^c$, and $(-\infty, a] = (a, \infty)^c$, so both $(-\infty, a)$ and $(-\infty, a]$ are measurable. Of course $(-\infty, \infty)$ is measurable.

Now suppose $J$ is bounded interval with two ends $a \leq b$. By Lemma $(a, b]$ is measurable, and we have shown $\{a\}$ is measurable. Since $[a, b] = \{a\} \cup (a, b]$, $(a, b) = \bigcup_{n=1}^{\infty}(a, b - \frac{1}{n}]$, so open / closed intervals are measurable. Since $[a, b) = \{a\} \cup (a, b)$ so $[a, b)$ is measurable too. Therefore any bounded interval is measurable.

2) *Any open subset of $\mathbb{R}$ is measurable, so is any closed subset of $\mathbb{R}$. Hence any closed / open subset is measurable.*

According to Question 7, Problem Sheet 1, if $G \subset \mathbb{R}$ is open, then $G$ has a decomposition

$$G = \bigcup_i (a_i, b_i) \tag{4.13}$$

where $(a_i, b_i)$ are *disjoint* open intervals (bounded or unbounded), *at most countably many.* Since $\mathcal{M}_{\text{Leb}}$ is $\sigma$-algebra, and each interval $(a_i, b_i)$ is measurable, hence $G$ is also measurable. Since $m^*$ is a measure on $\mathcal{M}_{\text{Leb}}$, so it is countably additive,

$$m^*(G) = \sum_i (b_i - a_i).$$

*Remark.* However we should note that it is in general impossible to arrange intervals $(a_i, b_i)$ in the decomposition of $G$ in an order, such that $a_1 < b_1 < a_2 < b_2 < \cdots$.

The Borel $\sigma$-algebra $\mathcal{B}(\mathbb{R})$ over $\mathbb{R}$ is defined to be the smallest $\sigma$-algebra containing all open subsets of $\mathbb{R}$. A subset $E \subset \mathbb{R}$ is called Borel measurable if $E \in \mathcal{B}(\mathbb{R})$. Since any open subset belongs to $\mathcal{M}_{\text{Leb}}$ and $\mathcal{M}_{\text{Leb}}$ is a $\sigma$-algebra, therefore $\mathcal{M}_{\text{Leb}}$ is bigger than $\mathcal{B}(\mathbb{R})$. We thus have the following important conclusion.

**Lemma 4.8** *$\mathcal{B}(\mathbb{R}) \subset \mathcal{M}_{Leb}$, that is, any Borel measurable subset is Lebesgue measurable.*

On the other hand, $\mathcal{M}_{\text{Leb}}$ is not lot bigger than $\mathcal{B}(\mathbb{R})$. In fact we have

**Lemma 4.9** *If $E \in \mathcal{M}_{Leb}$, then there are $A, B \in \mathcal{B}(\mathbb{R})$, such that $A \subset E \subset B$ and $m^*(E \setminus A) = m^*(B \setminus E) = 0$. That is, $\mathcal{M}_{Leb}$ and $\mathcal{B}(\mathbb{R})$ differ by Lebesgue measure zero subsets.*

*Example 1.* If $a \in \mathbb{R}$, then $m^*(\{a\}) = 0$ (see above), so that, by the countable additivity of $m^*$ on $\mathcal{M}_{\text{Leb}}$, any countable subset is measurable and has Lebesgue measure zero. In particular $m^*(\mathbb{Q}) = 0$.

*Example 2.* There are uncountable Lebesgue null sets. Cantor set (see pages 15 -17 in Lecture Notes) is an example. Read the notes carefully for its construction. Also Question 2 part (a), Problem Sheet 2.

*Example 3.* There are subsets which are not Lebesgue measurable. See page 17 Lecture Notes for an example.

To construct the example, we need the following fact: the Lebesgue outer measure $m^*$ (and therefore the Lebesgue measure $m^*$ on $\mathcal{M}_{\text{Leb}}$) is *translation invariant* in the sense that:

$$m^*(E + a) = m^*(E)$$

for every subset $E$, and for every number $a \in \mathbb{R}$, where $E + a = \{x + a : x \in E\}$. This is a part of Question 1 is Problem Sheet 2. Of course, by symmetry, you only need to show that $m^*(E + a) \leq m^*(E)$, which you can prove it by using the definition of $m^*$ given by (4.12).

Let us point out that examples of non-Lebesgue measurable subsets are not examinable.

Let me finish these notes by saying that there are Lebesgue measurable sets which are not Borel measurable, that is, the $\sigma$-algebra $\mathcal{M}_{\text{Leb}}$ is *strictly larger* than the Borel $\sigma$-algebra $\mathcal{B}(\mathbb{R})$.

## 4.3 The Cantor set and null subsets

**Example 4.10** *If $A$ is a countable subset, then $A$ is a Lebesgue null set.*

There are null sets which are not countable! Here is an example.

**Example 4.11** *The Cantor ternary set.*

Consider the closed interval $J_0^{(1)} = [0, 1]$.

1) Divide $J_0^{(1)}$ equally into three sub-intervals, the middle open interval $I_1^{(1)} = (\frac{1}{3}, \frac{2}{3})$ is removed from $J_0^{(1)}$, the remaining two closed sub-intervals are $J_1^{(1)} = [0, \frac{1}{3}]$ and $J_2^{(1)} = \left[\frac{2}{3}, 1\right]$, each has length $\frac{1}{3}$.

2) Repeat step 1) for each closed interval $J_1^{(1)}$ and $J_2^{(1)}$: divide each equally into three sub-intervals and remove the middle open ones. From $J_1^{(1)}$ we remove $I_1^{(2)} = \left(\frac{1}{3^2}, \frac{2}{3^2}\right)$, and from $J_2^{(1)}$ remove

$$I_2^{(2)} = \left(\frac{2}{3} + \frac{1}{3^2}, \frac{2}{3} + \frac{2}{3^2}\right) = \frac{2}{3} + \left(\frac{1}{3^2}, \frac{2}{3^2}\right).$$

The remaining $2^2$ closed intervals are denoted by $J_1^{(2)}, J_2^{(2)}, J_3^{(2)}, J_4^{(2)}$ each has length $\frac{1}{3^2}$.

3) Repeat the procedure for each $J_i^{(2)}$ , remove the middle open intervals $I_i^{(2)}$ $(i = 1, \cdots, 2^2)$ and the remaining $2^3$ closed intervals are denoted by $J_i^{(3)}$ for $i = 1, \cdots, 2^3$, each has length $\frac{1}{3^3}$.

Repeating this process, for each $n$, we have $2^{n-1}$ disjoint open intervals $I_i^{(n)}$ $(i = 1, \cdots, 2^{n-1})$ with equal length $\frac{1}{3^n}$, and $2^n$ disjoint closed intervals $J_i^{(n)}$ $(i = 1, 2, \cdots, 2^n)$ with length $\frac{1}{3^n}$(where $n = 1, 2, \cdots$).

For each $n$, $I_i^{(n)}$, $J_k^{(n)}$ $(i = 1, \cdots, 2^{n-1}$ and $k = 1, \cdots, 2^n)$ are disjoint sub-intervals of $[0, 1]$, and

$$\left(\bigcup_{k=1}^{n} \bigcup_{i=1}^{2^{k-1}} I_i^{(k)}\right) \bigcup \left(\bigcup_{k=1}^{2^n} J_k^{(n)}\right) = [0, 1].$$

Cantor's ternary set $C$ is defined to be the subset of $[0, 1]$ by removing all middle intervals $I_i^{(n)}$. That is

$$C = [0, 1] \setminus \left(\bigcup_{n=1}^{\infty} \bigcup_{i=1}^{2^{n-1}} I_i^{(n)}\right).$$

**Lemma 4.12** *The Cantor ternary set $C$ is a null set.*

**Proof.** By the construction, for every $n$, $J_i^{(n)}$ $(i = 1, \cdots, 2^n)$ is a finite cover of $C$, so that

$$m^*(C) \le \sum_{i=1}^{2^n} |J_i^{(n)}| = 2^n \frac{1}{3^n} \to 0$$

as $n \to \infty$. Therefore $m^*(C) = 0$, and $C$ is a null subset by definition. ■

*Ternary expansions* – For decimal expansion of $x = 0.x_1x_2 \cdots \in (0, 1]$ we mean that

$$x = \sum_{n=1}^{\infty} \frac{x_n}{10^n}.$$

Similarly $x \in (0, 1]$ can be written in its ternary expansion

$$x = \sum_{n=1}^{\infty} \frac{a_n}{3^n} = 0.a_1a_2 \cdots$$

where $a_n = 0, 1$ or 3, and we use the convention that it should not end with $a_k = 2$ for all $k \ge N$. Then, in terms of ternary expansion

$$(\frac{1}{3}, \frac{2}{3}) = (0, 1, 0.2)$$

and $x \in I_1^{(1)} = (\frac{1}{3}, \frac{2}{3})$ then $x = 0.a_1a_2 \cdots$ with $a_1 = 1$. Let $A = \{0.a_1a_2 \cdots :$ where $a_i = 0$ or $2\}$. Then by our construction, $A \subseteq C$ so that $m^*(A) \le m^*(C) = 0$.

We leave the reader as an exercise to show that 1) $C$ is closed, 2) $C$ is uncountable, and 3) describe $C$ in terms of ternary expansions for real numbers.

Finally we should point out that $\mathcal{M}_{\text{Bor}}$ is strictly smaller than $\mathcal{M}_{\text{Leb}}$. However we have the following fact

**Proposition 4.13** *If $E \in \mathcal{M}_{Leb}$ then there are $A, B \in \mathcal{M}_{Bor}$ such that $A \subset E \subset B$ and both $E \setminus A$ and $B \setminus E$ are null sets.*

The proof is outside the syllabus.

## 4.4 An example of non measurable sets

While not every subset of $\mathbb{R}$ is measurable, though it is not easy to produce an example of non measurable sets, which is a good news indeed. In fact we need to invoke *the axiom of choice* to produce a subset which is not in $\mathcal{M}_{\mathrm{Leb}}$.

If $E$ is a subset of $\mathbb{R}$, and $a$ is a real number, then $a+A=\{a+x : x \in A\}$ is a translation of a subset $A$. It is elementary that if $J$ is an interval, so is $a+J$ and $|a+J|=|J|$. According to the definition of the outer measure $m^*$ we may deduce that $m^*(a+A)=m^*(A)$. It follows that, if $E \in \mathcal{M}_{\mathrm{Leb}}$, then $a+E \in \mathcal{M}_{\mathrm{Leb}}$ and $m(a+E)=m(E)$. Therefore the Lebesgue measure $m$ is invariant under translations.

Consider the unit interval $[0,1]$. Define the following equivalence relations: if $x, y \in [0,1]$ and $x-y \in \mathbb{Q}$, then we say $x \sim y$. Divide $[0,1]$ into equivalent classes, and choose exactly one number from each equivalent class to form a subset $A \subset [0,1]$. We claim that $A$ is not measurable, i.e. $A \notin \mathcal{M}_{\mathrm{Leb}}$. In fact, list the rational numbers in $[-1,1]$ by $r_1, r_2, \cdots$. Then $\cup_{i=1}^{\infty}(r_i+A) \supseteq [0,1]$ and $\cup_{i=1}^{\infty}(r_i+A) \subset [-1,2]$, so that

$$1 \leq m^*\left(\bigcup_{i=1}^{\infty}(r_i+A)\right) \leq 3.$$

Note that $r_i+A$ are disjoint, and $m^*(r_i+A)=m^*(A)$ for every $i$. If $A$ were measurable, then $r_i+A$ are measurable for every $i$, hence, by countable additivity, we would have

$$m^*\left(\bigcup_{i=1}^{\infty}(r_i+A)\right)=\sum_{i=1}^{\infty} m^*(r_i+A)=\sum_{i=1}^{\infty} m^*(A)$$

so that

$$1 \leq \sum_{i=1}^{\infty} m^*(A) \leq 3$$

which is impossible as $\sum_{i=1}^{\infty} m^*(A)=0$ or $\sum_{i=1}^{\infty} m^*(A)=\infty$ (according to $m^*(A)=0$ or not).

Thus $A$ is not Lebesgue measurable.

## 4.5 Lebesgue sub-spaces

Suppose $E \subset \mathbb{R}$ is Lebesgue measurable, then

$$\begin{aligned}
\mathcal{M}_{\mathrm{Leb}}(E) &= \{E \cap A : \text{where } A \in \mathcal{M}_{\mathrm{Leb}}\} \\
&= \{A : \text{where } A \subseteq E \text{ and } A \in \mathcal{M}_{\mathrm{Leb}}\}
\end{aligned}$$

is a $\sigma$-algebra on $E$. On the other hand $\mathcal{M}_{\mathrm{Leb}}(E) \subseteq \mathcal{M}_{\mathrm{Leb}}$ so that the restriction of the Lebesgue measure $m$, denoted by $m|_E$ or by $m$ if no confusion is possible, is obviously a measure on $(E, \mathcal{M}_{\mathrm{Leb}}(E))$, called the Lebesgue measure on $E$. The triple $(E, \mathcal{M}_{\mathrm{Leb}}(E), m)$ is therefore a measure space, called a Lebesgue subspace. A function $f$ defined on $E$ which is $\mathcal{M}_{\mathrm{Leb}}(E)$-measurable is called Lebesgue measurable on $E$. $\mathcal{M}_{\mathrm{Leb}}(E)$ will be simply denoted (by abusing notations) by $\mathcal{M}_{\mathrm{Leb}}$ if no confusion may arise.

# 5 Measurable functions

In this section we identify a class of functions, called measurable functions, for which we aim to define integrals. The concept of measurability can be developed independent of a measure, we therefore develop this concept for a general measurable space $(\Omega, \mathcal{F})$.

## 5.1 Definition and basic properties

Recall that $\mathcal{B}(\mathbb{R})$ (or denoted by $\mathcal{M}_{\mathrm{Bor}}$) denotes the Borel $\sigma$-algebra on $\mathbb{R}$, which is the smallest $\sigma$-algebra containing all open sets in $\mathbb{R}$, or equivalently, the smallest $\sigma$-algebra which contains all intervals.

Let us begin with some comments on notations. If $f$ is a function defined on a space $\Omega$ and takes its values in $[-\infty, \infty]$ (i.e. $f$ is a mapping from $\Omega$ to $[-\infty, \infty]$), and if $A$ is a subset of $[-\infty, \infty]$, then $f^{-1}(A)$ is the pre-image of $A$ under $f$, that is,

$$f^{-1}(A) = \{x \in \Omega : f(x) \in A\}.$$

For simplicity, we will also use $\{f \in A\}$ to denote $f^{-1}(A)$. More generally, if $P$ is a property (or a statement) depending on $x \in \Omega$, then we will use $\{P\}$ to denote the subset $\{x \in \Omega : P(x) \text{ holds }\}$, if the underlying space $\Omega$ is clear, and if no confusion may arise. For example, $f$ as above and $a \in \mathbb{R}$, then $\{f > a\} = \{x \in \Omega : f(x) > a\}$. That is, $\{f > a\}$ is the pre-image of $(a, \infty]$ under $f$. Similarly $\{f = \infty\}$ denotes the set $\{x \in \Omega : f(x) = \infty\}$ etc. As another example, if $f$ and $g$ are two mappings from $\Omega$ into $[-\infty, \infty]$, then

$$\{f \neq g\} = \{x \in \Omega : f(x) = g(x)\}$$

and

$$\{f > g\} = \{x \in \Omega : f(x) > g(x)\}$$

and etc.

Let $\mathcal{F}$ be a $\sigma$-algebra on the space $\Omega$.

**Definition 5.1** *Let $f : \Omega \to \mathbb{R}$ be a function on $\Omega$. Then $f$ is $\mathcal{F}$-measurable if $f^{-1}(G) \in \mathcal{F}$ for every $G \in \mathcal{B}(\mathbb{R})$. That is, $\{f \in G\} \in \mathcal{F}$ for every Borel subset $G$.*

In probability theory, an $\mathcal{F}$-measurable function on $\Omega$ is called a random variable on the measurable space $(\Omega, \mathcal{F})$. To discuss the measurability of functions (in general mappings, the following fact is very useful.

**Lemma 5.2** *Let $X : \Omega \to S$ be a mapping, and $\mathcal{F}$ be a $\sigma$-algebra on $\Omega$. Then*

$$\mathcal{F}_X = \left\{A \subset S : X^{-1}(A) \in \mathcal{F}\right\} \tag{5.1}$$

*is a $\sigma$-algebra. $\mathcal{F}_X$ is the push-forward $\sigma$-algebra of $\mathcal{F}$ by $X$.*

The proof is left as an exercise [Problem 4, Sheet 2].

**Proposition 5.3** *Suppose $f : \Omega \to \mathbb{R}$. Then the following statements are equivalent:*

*1) $f^{-1}(G) \in \mathcal{F}$ for every $G \in \mathcal{M}_{Bor}$, i.e. $f$ is $\mathcal{F}$-measurable.*

*2) $\{f > a\} \in \mathcal{F}$ for every $a \in \mathbb{R}$,*

*3) $\{f < a\} \in \mathcal{F}$ for every $a \in \mathbb{R}$,*

*4) $\{f \leq a\} \in \mathcal{F}$ for every $a \in \mathbb{R}$,*

*5) $f^{-1}(J) \in \mathcal{F}$ for every interval $J$.*

**Proof.** Let $\mathcal{G} = \mathcal{F}_f$ be defined by (5.1). Then $\mathcal{G}$ is a $\sigma$-algebra, and $f$ is $\mathcal{F}$-measurable, i.e. it satisfies 1), if and only if $\mathcal{B}(\mathbb{R}) \subset \mathcal{G}$. Let us for example show that 1) and 4) are equivalently. Clearly 1) implies 4), so let us prove the other direction that 4) yields 1). To this end, let $\mathscr{A}$ be the family of all intervals $(-\infty, a]$. Then 4) says that $\mathscr{A} \subset \mathcal{G}$. Since $\mathcal{G}$ is a $\sigma$-algebra, so that $\sigma\{\mathscr{A}\} \subset \mathcal{G}$, where $\sigma\{\mathscr{A}\}$ is the smallest $\sigma$-algebra which contains all subsets in $\mathscr{A}$. Since $\sigma\{\mathscr{A}\} = \mathcal{B}(\mathbb{R})$, therefore $\mathcal{B}(\mathbb{R}) \subset \mathcal{G}$. Thus for every $G \in \mathcal{B}(\mathbb{R})$, $\{f \in G\} \in \mathcal{F}$. ■

**Definition 5.4** *Let $\Omega$ be a Lebesgue measurable subset of $\mathbb{R}$, and $\mathcal{M}_{Leb}(\Omega)$ (or $\mathcal{M}_{Leb}$ if no confusion may arise) denote the $\sigma$-algebra of all Lebesgue measurable subsets of $\Omega$, that is,*

$$\mathcal{M}_{Leb}(\Omega) = \{A \in \mathcal{M}_{Leb} : A \subset \Omega\}.$$

*Then $(\Omega, \mathcal{M}_{Leb}(\Omega))$ is a measurable space, and the Lebesgue measure restricted on $\mathcal{M}_{Leb}(\Omega)$ is a measure on $(\Omega, \mathcal{M}_{Leb})$. The measure space $(\Omega, \mathcal{M}_{Leb}, m)$ is called a Lebesgue measure space. Let $f : \Omega \to \mathbb{R}$ be a function defined on a Lebesgue measurable subset $\Omega$.*

*If $f : \Omega \to \mathbb{R}$ is $\mathcal{M}_{Leb}$-measurable, then $f$ is called Lebesgue measurable on $\Omega$.*

**Definition 5.5** *A function $f : \mathbb{R} \to \mathbb{R}$ which is $\mathcal{B}(\mathbb{R})$-measurable, is called* Borel measurable.

Since $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}_{\text{Leb}}$, a Borel measurable function is Lebesgue measurable. The converse is not true in general, there are Lebesgue measurable functions which are not Borel measurable.

*Example.* 1) Suppose $h : \mathbb{R} \to \mathbb{R}$ is continuous. Then if $U$ is open then $h^{-1}(U)$ is open, so that $h^{-1}(U) \in \mathcal{B}(\mathbb{R})$. Since

$$\mathcal{G} = \left\{A \subset \mathbb{R} : h^{-1}(A) \in \mathcal{B}(\mathbb{R})\right\}$$

is a $\sigma$-algebra, and any open subset belongs to $\mathcal{G}$, by definition, $\mathcal{B}(\mathbb{R}) \subset \mathcal{G}$. Therefore $h^{-1}(G) \in \mathcal{B}(\mathbb{R})$ for any $G \in \mathcal{B}(\mathbb{R})$. Thus a continuous function $h$ is $\mathcal{B}(\mathbb{R})$-measurable.

2) If $h : (a, b) \to \mathbb{R}$ is monotone, then $h$ is Borel measurable.

**Proposition 5.6** *Let $(\Omega, \mathcal{F})$ be a measurable space.*

*1) Let $A \subset \Omega$. Then $1_A$ is $\mathcal{F}$-measurable if and only if $A \in \mathcal{F}$.*

*2) If $f : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable, and $h : \mathbb{R} \to \mathbb{R}$ is Borel measurable, then $h \circ f$ is $\mathcal{F}$-measurable. In particular, if $h$ is continuous, and $f$ is $\mathcal{F}$-measurable, then $h \circ f$ is also $\mathcal{F}$-measurable. [*While, on the other hand, there are Lebesgue measurable functions $h$ and $f$ such that $h \circ f$ is not Lebesgue measurable. *]*

*3) If $a \in \mathbb{R}$ and $f, g : \Omega \to \mathbb{R}$ are $\mathcal{F}$-measurable, then $af$, $f \pm g$, $fg$, $f/g$ (if $g \neq 0$) are $\mathcal{F}$-measurable.*

*4) $f \wedge g = \max\{f, g\}$ and $f \vee g = \min\{f, g\}$ are $\mathcal{F}$-measurable. Hence $|f|$, $f^+$ and $f^-$ are $\mathcal{F}$-measurable, where $f^+ = f \vee 0$ and $f^- = (-f) \vee 0$. $f^+$ (resp. $f^-$) is called the positive part (negative part) of $f$. $|f| = f^+ + f^-$ and $f = f^+ - f^-$.*

**Proof.** 1) For every $G$, $\{1_A \in G\}$ equals $A$, $A^c$, $\varnothing$ or $\Omega$, so $I_A$ is measurable if and only if $A \in \mathcal{F}$.

2) In fact, for any $G \in \mathcal{B}(\mathbb{R})$, $h^{-1}(G) \in \mathcal{B}(\mathbb{R})$, thus $(h \circ f)^{-1}(G) = f^{-1}(h^{-1}(G)) \in \mathcal{F}$, so $h \circ f$ is $\mathcal{F}$-measurable.

3) There is nothing to prove if $a = 0$. If $a > 0$ then for any $b$ we have

$$\{af > b\} = \left\{ f > \frac{b}{a} \right\} \in \mathcal{F}$$

and if $a < 0$ then

$$\{af > b\} = \left\{ f < \frac{b}{a} \right\} \in \mathcal{F}$$

so that $af$ is measurable. In particular $-g$ is measurable if $g$ is measurable. We show that $f + g$ is measurable. For any $b$ we have

$$\begin{aligned} \{f + g \quad > \quad b\} &= \bigcup_{q \in \mathbb{Q}} \{f > q \text{ and } g > b - q\} \\ &= \bigcup_{q \in \mathbb{Q}} \{f > q\} \cap \{g > b - q\}. \end{aligned}$$

Since $f, g$ are measurable so that $\{f > q\} \in \mathcal{F}$ and $\{g > b - q\} \in \mathcal{F}$ and therefore

$$\{f > q\} \cap \{g > b - q\} \in \mathcal{F} \quad \forall q.$$

Since the set $\mathbb{Q}$ of all rational numbers is countable, it follows that

$$\bigcup_{q \in \mathbb{Q}} \{f > q\} \cap \{g > b - q\} \in \mathcal{F}$$

and therefore $f + g$ is measurable.

Since $f^2 = h \circ f$ where $h(x) = x^2$ is a continuous function, so that $f^2$ is measurable if $f$ is measurable. Now

$$fg = \frac{1}{4}\left[(f + g)^2 - (f - g)^2\right]$$

is measurable.

Similarly, since $h(x) = |x|$ is continuous, so that $|f| = h \circ f$ is measurable. Hence

$$f \vee g = \frac{1}{2}\left[(f + g) + |f - g|\right]$$

and

$$f \wedge g = \frac{1}{2}\left[(f + g) - |f - g|\right]$$

are measurable. In particular $f^+ = f \vee 0$ and $f^- = (-f) \vee 0$ are measurable. ∎

**Definition 5.7** *A function $f : \Omega \to [-\infty, \infty]$ is $\mathcal{F}$-measurable, if $\{f = -\infty\}$, $\{f = \infty\}$ are measurable, and $f^{-1}(G) \in \mathcal{F}$ for every Borel subset $G$.*

Next we prove that the class of $\mathcal{F}$-measurable functions on a measurable space $(\Omega, \mathcal{F})$ is closed under limiting operations.

**Proposition 5.8** *If $\{f_n : n = 1, 2, \cdots\}$ is a sequence of $\mathcal{F}$-measurable functions, then $\sup_{n\geq 1} f_n$, $\inf_{n\geq 1} f_n$, $\limsup_{n\to\infty} f_n$ and $\liminf_{n\to\infty} f_n$ are $\mathcal{F}$-measurable. In particular, if $f = \lim_{n\to\infty} f_n$ exists, then $f$ is $\mathcal{F}$-measurable.*

**Proof.** For each $n$ consider

$$g_n(x) = \sup\{f_n(x), f_{n+1}(x), \cdots\}$$

and

$$h_n(x) = \inf\{f_n(x), f_{n+1}(x), \cdots\}.$$

Then $\{g_n : n = 1, 2, \cdots\}$ is a [point-wise] decreasing sequence and $\{h_n : n = 1, 2, \cdots\}$ is an increasing sequence, hence $g_n \downarrow g$ and $h_n \uparrow h$, where $g$ and $h$ may take value $\pm\infty$. $g$ and $h$ are called the upper limit and lower limit of $(f_n)$, denoted by $\overline{\lim}_{n\to\infty} f_n$ (or $\limsup f_n$) and $\underline{\lim}_{n\to\infty} f_n$ (or by $\liminf f_n$) respectively.

For each $n$, $g_n$ takes values in $(-\infty, \infty]$. For any real number $a$ we have

$$\{g_n \leq a\} = \bigcap_{m=n}^{\infty} \{f_m \leq a\}$$

and

$$\{g_n = \infty\} = \left(\bigcup_{N=1}^{\infty} \{g_n \leq N\}\right)^c$$

so that $g_n$ is $\mathcal{F}$-measurable. On the other hand

$$h_n = -\inf\{-f_n(x), -f_{n+1}(x), \cdots\}$$

so that $h_n$ is $\mathcal{F}$-measurable. Finally since

$$\limsup_{n\to\infty} f_n = \inf\{g_n : n \geq 1\}$$

thus $\limsup_{n\to\infty} f_n$ is $\mathcal{F}$-measurable, and $\underline{\lim}_{n\to\infty} f_n = -\overline{\lim}_{n\to\infty}(-f_n)$ is also $\mathcal{F}$-measurable. ∎

If $(\Omega, \mathcal{F})$ is a measurable space, then a function $\varphi : \Omega \to \mathbb{R}$ is called a *simple function*, or more precisely called a *simple $\mathcal{F}$-measurable function* on $\Omega$, if

$$\varphi = \sum_{i=1}^{k} c_i 1_{E_i}$$

for some positive integer $k$, some reals $c_i$ and some $\mathcal{F}$-measurable subsets $E_i$.

The collection of all *non-negative, simple* functions is denoted by $\mathcal{S}^+(\Omega, \mathcal{F})$ or by $\mathcal{S}^+$ if no confusion may arise.

**Theorem 5.9** *Suppose that $f$ is a measurable function taking values in $[0, \infty]$. Then there is an increasing sequence of simple functions $(f_n)$ such that $f_n \uparrow f$.*

**Proof.** For each $n$ we define

$$f_n = \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} 1_{E_k^{(n)}} + 2^n 1_{A_n},$$

where

$$E_k^{(n)} = \left\{ x : \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \right\} \text{ and } A_n = \{x : f(x) \geq 2^n\}$$

are measurable. Then $(f_n)$ is an increasing sequence of non-negative simple $\mathcal{F}$-measurable functions. Moreover $0 \leq f - f_n \leq \frac{1}{2^n}$ on $\{f < 2^n\}$ and $f_n = 2^n$ on $\{f \geq 2^n\}$. Therefore $f_n \uparrow f$ as $n \to \infty$. ■

Therefore we have the following simple structure theorem for measurable functions in terms of simple measurable functions.

**Corollary 5.10** *A function $f$ on $\Omega$ taking values in $[-\infty, \infty]$ is $\mathcal{F}$-measurable if and only if there is a sequence of simple $\mathcal{F}$-measurable functions $f_n : \Omega \to \mathbb{R}$ such that $f_n \to f$.*

**Proof.** We note that $f = f^+ - f^-$ is measurable if and only if both $f^+$ and $f^-$ are measurable. Apply the previous theorem to $f^+$ and $f^-$. ■

## 5.2 Almost everywhere properties

Let $(\Omega, \mathcal{F}, \mu)$ be a measurable space. If $P$ is a property depending on $x \in \Omega$, then $\{P\}$ denotes the subset $\{x \in \Omega : P(x) \text{ holds }\}$, and $\mu[P]$ deontes the measure of $\{P\}$ if $\{P\}$ is measurable. That is,

$$\mu[P] = \mu(\{P\}) = \mu(\{x \in \Omega : P(x) \text{ holds }\}).$$

For example, if $f : \Omega \to [-\infty, \infty]$ is $\mathcal{F}$-measurable and $\lambda$ is a real number, then

$$\mu[f > \lambda] = \mu(\{x \in \Omega : f(x) > \lambda\}).$$

If $f$ and $g$ are two measurable functions, then

$$\mu[f \neq g] = \mu(\{x \in \Omega : f(x) \neq g(x)\})$$

etc.

We say the property $P$ holds $\mu$-*almost everywhere* (or almost surely) on $\Omega$, if

$$\mu[P \text{ doesn't hold }] = \mu(\{x \in \Omega : \; P(x) \text{ doesn't hold }\}) = 0.$$

If the underlying space $\Omega$ and the measure $\mu$ is clear from the context, then we simply say the property $P$ holds almost everywhere. For example, if $f$ and $g$ are two functions, then we say $f = g$ almost everywhere if $\mu[f \neq g] = 0$.

If $(f_n)$ is a sequence of functions, then $f_n$ converges to $f$ almost everywhere, if

$$\mu(\{x \in \Omega : f_n(x) \text{ does not converge to } f(x)\}) = 0.$$

Recall that a measure space $(\Omega, \mathcal{F}, \mu)$ is complete, if $A \in \mathcal{F}$ and $\mu(A) = 0$, then any subset of $A$ is $\mathcal{F}$-measurable, that is, any subset of $A$ also belongs to $\mathcal{F}$. Any Lebesgue measurable sub-space $(E, \mathcal{M}_{\text{Leb}}, m)$ (where $E$ is Lebesgue measurable) is complete. However, $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$ is not complete.

**Proposition 5.11** *Suppose $(\Omega, \mathcal{F}, \mu)$ is a* complete *measure space. If $f : \Omega \to [-\infty, \infty]$ is $\mathcal{F}$-measurable, and $g = f$ almost everywhere, then $g$ is also $\mathcal{F}$-measurable.*

This is an exercise in Problem Sheet 2.

**Proposition 5.12** *Suppose $(\Omega, \mathcal{F}, \mu)$ is a* complete *measure space. If $f_n : \Omega \to [-\infty, \infty]$ are $\mathcal{F}$-measurable functions, and $f_n \to f$ almost everywhere on $\Omega$, then $f$ is $\mathcal{F}$-measurable.*

## 5.3 Examples

We give in this part further properties about measurable functions in terms of examples.

*Example 1.* If $g : \mathbb{R} \to \mathbb{R}$ is monotone, then $g$ is Borel measurable. In fact if $I$ is an interval, then $g^{-1}(I)$ is a union of at most countable many intervals, so it belongs to $\mathcal{B}(\mathbb{R})$. Hence $g$ is Borel measurable.

*Example 2.* Let $(\Omega, \mathcal{F})$ be a measurable space. Suppose $f_n : \Omega \to \mathbb{R}$ are $\mathcal{F}$-measurable. Recall that $\{f_n\}$ converges at $x \in \Omega$ if and only if $\{f_n(x)\}$ is a Cauchy sequence. Hence

$$\begin{aligned}&\{x \in \Omega : (f_n(x)) \text{ converges to a number}\}\\ =&\bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{m,n=N}^{\infty} \left\{ |f_n - f_m| < \frac{1}{k} \right\}\end{aligned}$$

and

$$\begin{aligned}&\{x \in \Omega : (f_n(x)) \text{ doesn't converge to a number}\}\\ =&\bigcup_{k=1}^{\infty} \bigcap_{N=1}^{\infty} \bigcup_{m,n=N}^{\infty} \left\{ |f_n - f_m| \geq \frac{1}{k} \right\}\end{aligned}$$

are both measurable.

*Example 3.* Let $(\Omega, \mathcal{F})$ be a measurable space. If $f_n, f : \Omega \to \mathbb{R}$ are $\mathcal{F}$-measurable, then

$$\{f_n \to f\} = \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \left\{ |f_n - f| < \frac{1}{k} \right\}$$

and

$$\{f_n \nrightarrow f\} = \bigcup_{k=1}^{\infty} \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} \left\{ |f_n - f| \geq \frac{1}{k} \right\}$$

are measurable as well.

*Example 4.* Let $E \subset \mathbb{R}$, and $f, f_n : E \to \mathbb{R}$ $(n = 1, 2, \cdots)$ be Lebesgue measurable functions on $E$. Then $f_n \to f$ almost everywhere on $E$ (with respect to the Lebesgue measure), i.e.

$$m\,[f_n \text{ doesn't converge to } f] = 0$$

[where
$$\{f_n \text{ doesn't converge to } f\} = \{x \in E : f_n(x) \nrightarrow f(x)\}$$
and similar notations apply to other sets], if and only if for every $\varepsilon > 0$
$$m\left(\bigcap_{N=1}^{\infty}\bigcup_{n=N}^{\infty}\{|f_n - f| \geq \varepsilon\}\right) = 0 \ . \tag{5.2}$$

If in addition $m(E) < \infty$, then [Proposition 2.4, item 3)] (5.2) is equivalent to the condition that
$$\lim_{N\to\infty} m\left(\bigcup_{n=N}^{\infty}\{|f_n - f| \geq \varepsilon\}\right) = 0 \tag{5.3}$$
for any $\varepsilon > 0$.

*Example 5.* Let $E = (0, \infty)$, and $f_n(x) = 1$ if $x \in (0, n]$ and $f_n(x) = 0$ if $x > n$. Then $f_n \to f = 1_{(0,\infty)}$ everywhere. For every $\varepsilon \in (0, 1)$ we have
$$\{|f_n - f| \geq \varepsilon\} = (n, \infty)$$
so that
$$m\left(\{|f_n - f| \geq \varepsilon\}\right) = \infty.$$
Thus (5.3) doesn't hold. On the other hand
$$\bigcap_{N=1}^{\infty}\bigcup_{n=N}^{\infty}\{|f_n - f| \geq \varepsilon\} = \bigcap_{N=1}^{\infty}(N, \infty) = \emptyset.$$

[This is a simple example that $f_n \to f$ everywhere, but $f_n$ does not converge to $f$ *in measure*, see Definition 7.5 below].

*Example 6.* (*Egorov's theorem*) Let $E \subset \mathbb{R}$ be measurable such that $m(E) < \infty$. Suppose $f_n, f : E \to \mathbb{R}$ are measurable, and suppose $f_n \to f$ almost everywhere on $E$. Then for every $\delta > 0$ there is a measurable $E_\delta \subset E$ such that $m(E \setminus E_\delta) < \delta$ and $f_n \to f$ uniformly on $E_\delta$.

*Proof.* Since $f_n \to f$ almost surely on $E$ and $m(E) < \infty$, according to (5.3) in Example 4, for every $\delta > 0$ and for any $k = 1, 2, \cdots$ (applying (5.2) to $\varepsilon = \frac{1}{k}$) there is $n_k$ such that
$$m\left(\bigcup_{j=n}^{\infty}\left\{|f_j - f| \geq \frac{1}{k}\right\}\right) \leq \frac{\delta}{2^k} \text{ for } n \geq n_k.$$

Let
$$E_k = \bigcap_{j=n_k}^{\infty}\left\{|f_j - f| < \frac{1}{k}\right\} = \left\{|f_j - f| < \frac{1}{k} : j \geq n_k\right\}$$
and $E_\delta = \bigcap_{k=1}^{\infty} E_k$. Then $m(E_k^c) \leq \frac{\delta}{2^k}$ and
$$m(E \setminus E_\delta) = m\left(\bigcup_{k=1}^{\infty} E_k^c\right) \leq \sum_{k=1}^{\infty} m(E_k^c) \leq \sum_{k=1}^{\infty}\frac{\delta}{2^k} = \delta.$$

We claim that $f_n \to f$ uniformly on $E_\delta$. In fact, $x \in E_\delta$ if and only if $x \in E_k$ for all $k = 1, 2, \cdots$, and if and only if

$$|f_j(x) - f(x)| < \frac{1}{k} \quad \forall j \geq n_k.$$

Therefore

$$\sup_{x \in E_\delta} |f_j(x) - f(x)| \leq \frac{1}{k}, \;\; \forall j \geq n_k$$

and $f_n \to f$ uniformly on $E_\delta$.

# 6 Lebesgue integration

In this section we study Lebesgue's theory of integration. Lebesgue's theory of integration can be built on a measure space, though most of textbooks on integration theory deal with the case of Lebesgue measure first, the process to construct the theory of integration however does not become simpler if we restricted to the Lebesgue measure.

First of all we need to define the class of simple (measurable) functions, and we define integrals for simple and non-negative functions, then define integrals for non-negative measurable functions. Finally we define Lebesgue integrals for measurable functions.

In this course, we only deal with the theory of Lebesgue's integration for the Lebesgue measure, we will nevertheless develop the theory on a measure space $(\Omega, \mathcal{F}, \mu)$ if no additional effort is required and if only modifications of notations are required. We will however concentrate on the Lebesgue space $(E, \mathcal{M}_{\mathrm{Leb}}, m)$, where $E \subset \mathbb{R}$ is a Lebesgue measurable subset, and $m$ is the Lebesgue measure as our model of the measure space, and you may read notes as if $\Omega = E$, $\mathcal{F} = \mathcal{M}_{\mathrm{Leb}}(E)$, $\mu = m$ for some Lebesgue measurable subset $E \subset \mathbb{R}$, unless otherwise specified, and thus $\mathcal{F}$-measurable means Lebesgue measurable on $E$.

## 6.1 Lebesgue integrals, and integrable functions

Therefore we will develop Lebesgue's theory on a measure space $(\Omega, \mathcal{F}, \mu)$. The following technical conditions are enforced in what follows, unless said otherwise.

1) $(\Omega, \mathcal{F}, \mu)$ is complete in the sense that: if $A \in \mathcal{F}$ and $\mu(A) = 0$, then any subset of $A$ belongs to $\mathcal{F}$.

2) $(\Omega, \mathcal{F}, \mu)$ is $\sigma$-finite, that is, there is an increasing sequence $G_n \in \mathcal{F}$ ($n = 1, 2, \cdots$), such that $\bigcup_{n=1}^{\infty} G_n = \Omega$ and $\mu(G_n) < \infty$ for all $n$.

Our basic example is a Lebesgue sub-space $(E, \mathcal{M}_{\mathrm{Leb}}, m)$, where $E \subseteq \mathbb{R}$ is a Lebesgue measurable subset, and $\mathcal{M}_{\mathrm{Leb}}$ denotes the $\sigma$-algebra $E \cap \mathcal{M}_{\mathrm{Leb}}$ on $E$ for simplicity.

$\mathcal{S}^+(\Omega, \mathcal{F})$ denotes the collection of all non-negative, simple, $\mathcal{F}$-measurable functions on $\Omega$.

Suppose $\varphi = \sum_{i=1}^{k} c_i 1_{E_i}$ is a non-negative, simple, $\mathcal{F}$-measurable function on $\Omega$, where $c_i \geq 0$ and $E_i \in \mathcal{F}$, then its Lebesgue integral is defined by

$$\int_\Omega \varphi \mathrm{d}\mu = \sum_{i=1}^{k} c_i \mu(E_i)$$

where the convention that $0 \cdot \infty = 0$ has been used, to ensure the finite sum on the right-hand side is well defined and is independent of the representation of $\varphi$. In fact, if

$$\varphi = \sum_{i=1}^{k} c_i 1_{E_i} = \sum_{j=1}^{N} a_j 1_{A_j}$$

where $c_i \geq 0$ and $a_j \geq 0$ for all $i, j$, then

$$\sum_{i=1}^{k} c_i \mu(E_i) = \sum_{j=1}^{N} a_j \mu(A_j)$$

which follows from the additivity of the measure $\mu$.

**Proposition 6.1** *Let $\varphi, \psi \in \mathcal{S}^+(\Omega, \mathcal{F})$, and $\lambda \geq 0$. Then*
*1) $\int_\Omega (\varphi + \psi) d\mu = \int_\Omega \varphi d\mu + \int_\Omega \psi d\mu$.*
*2) $\int_\Omega \lambda \varphi d\mu = \lambda \int_\Omega \varphi d\mu$.*
*3) If $\varphi \leq \psi$ then $\int_\Omega \varphi d\mu \leq \int_\Omega \psi d\mu$.*

**Proof.** We may choose measurable sets $E_i \in \mathcal{F}$ $(i = 1, \cdots, k)$ which are disjoint such that

$$\varphi = \sum_{i=1}^{k} c_i 1_{E_i}, \quad \psi = \sum_{i=1}^{k} d_i 1_{E_i}$$

so that

$$\varphi + \psi = \sum_{i=1}^{k} (c_i + d_i) 1_{E_i}$$

and therefore

$$\int_\Omega (\varphi + \psi) d\mu = \sum_{i=1}^{k} (c_i + d_i) \mu(E_i) = \int_\Omega \varphi d\mu + \int_\Omega \psi d\mu,$$

which proves 1). If $\varphi \leq \psi$ then $c_i \leq d_i$ for all $i$ for which $E_i \neq \emptyset$, then

$$\int_\Omega \varphi d\mu = \sum_{i=1}^{k} c_i \mu(E_i) \leq \sum_{i=1}^{k} d_i \mu(E_i) = \int_\Omega \psi d\mu$$

which is 3). ■

Suppose $f : \Omega \to [0, \infty]$ is $\mathcal{F}$-measurable, then the (Lebesgue) integral of $f$ is defined by

$$\int_\Omega f d\mu = \sup \left\{ \int_\Omega \varphi d\mu : \varphi \in \mathcal{S}^+(\Omega, \mathcal{F}) \text{ s.t. } \varphi \leq f \right\}. \tag{6.1}$$

A non-negative measurable function $f$ is *integrable* on $\Omega$ (with respect to the measure $\mu$) if $\int_\Omega f d\mu < \infty$.

By definition, if $f, g$ are non-negative and measurable, $f \leq g$, and $\lambda$ be a non-negative number, then

$$\int_\Omega \lambda f d\mu = \lambda \int_\Omega f d\mu \text{ and } \int_\Omega f d\mu \leq \int_\Omega g d\mu$$

In particular, if $g$ is integrable, then so is $f$.

**Lemma 6.2** *$f : \Omega \to [0, \infty]$ is $\mathcal{F}$-measurable, then*

$$\mu[f \geq \lambda] \leq \frac{1}{\lambda} \int_{\Omega} f d\mu \tag{6.2}$$

*for every $\lambda > 0$, which is called* the Markov inequality.

We recall here the convention of our notation: if $P$ is a statement depending on $x \in \Omega$, then $\{P\} = \{x \in \Omega : P(x)\}$ and

$$\mu[P] = \mu(\{P\}) = \mu(\{x \in \Omega : P(x)\}).$$

Therefore $\{f \geq \lambda\} = \{x \in \Omega : f(x) \geq \lambda\}$ and

$$\mu[f \geq \lambda] = \mu(\{x \in \Omega : f(x) > \lambda\}).$$

**Proof.** *of Lemma* 6.2. Since $\{f \geq \lambda\}$ is $\mathcal{F}$-measurable, and $f$ is non-negative, the simple function $\varphi = \lambda 1_{\{f \geq \lambda\}} \leq f$. By definition of integration,

$$\int_{\Omega} f d\mu \geq \int_{\Omega} \varphi d\mu = \lambda \mu[f \geq \lambda],$$

which yields (7.1). ■

**Proposition 6.3** *$f : \Omega \to [0, \infty]$ is* integrable*, then $\mu[f = \infty] = 0$. That is, $f$ is finite almost everywhere on $\Omega$.*

**Proof.** Since $\{f = \infty\}$ is measurable, so that, by the Markov inequality

$$\mu[f = \infty] \leq \mu[f > \lambda] \leq \frac{1}{\lambda} \int_{\Omega} f d\mu$$

for every $\lambda > 0$, where the first inequality follows from the fact that $\{f = \infty\} \subset \{f > \lambda\}$ for every $\lambda > 0$. Letting $\lambda \uparrow \infty$ in the inequality above, we deduce that

$$0 \leq \mu[f = \infty] \leq 0$$

and therefore we must have $\mu[f = \infty] = 0$. ■

**Proposition 6.4** *If $f : \Omega \to [0, \infty]$ is $\mathcal{F}$-measurable, and $\int_{\Omega} f d\mu = 0$, then $f = 0$ almost everywhere on $\Omega$, i.e. $\mu[f \neq 0] = 0$.*

**Proof.** By Markov's inequality

$$\mu[f \geq \lambda] \leq \frac{1}{\lambda} \int_{\Omega} f d\mu = 0,$$

so that $\mu[f \geq \lambda] = 0$ for every $\lambda > 0$. Since $\{f > 0\} = \bigcup_{n=1}^{\infty} \left\{f \geq \frac{1}{n}\right\}$, we have

$$0 \leq \mu[f > 0] \leq \sum_{n=1}^{\infty} \mu\left[f \geq \frac{1}{n}\right] = 0.$$

Therefore $f = 0$ almost everywhere on $\Omega$. ■

*Lebesgue integrable functions on* $(\Omega, \mathcal{F}, \mu)$

If $f : \Omega \to [-\infty, \infty]$ is $\mathcal{F}$-measurable, then its positive part $f^+ = \max\{f, 0\}$ and negative part $f^- = \max\{-f, 0\}$ are $\mathcal{F}$-measurable on $\Omega$, $f = f^+ - f^-$ and $|f| = f^+ + f^-$. Thus $\int_\Omega f^+ d\mu$ and $\int_\Omega f^- d\mu$ are both well defined (may be $\infty$). If both $\int_\Omega f^+ d\mu$ and $\int_\Omega f^- d\mu$ are finite, then we say $f$ is *(Lebesgue) integrable* on $\Omega$, and define its (Lebesgue) integral

$$\int_\Omega f d\mu = \int_\Omega f^+ d\mu - \int_\Omega f^- d\mu.$$

Let $L^1(\Omega, \mathcal{F}, \mu)$ ($L^1(\Omega, \mu)$ or $L^1(\Omega)$ if no confusion is possible from the context, for simplicity) denote the space of all integrable functions on $\Omega$.

If $f : \Omega \to [-\infty, \infty]$ is measurable, then $f \in L^1(\Omega)$ if and only if $\int_\Omega |f| d\mu < \infty$. If $f \in L^1(\Omega)$, then $\mu\left[|f| = \infty\right] = 0$ [Applying Proposition 6.3 to $|f|$], that is, $f$ is finite almost everywhere on $\Omega$.

We have a very simple *comparison theorem* for integrability.

**Proposition 6.5** *1) Let $f$ and $g$ be $\mathcal{F}$-measurable. If $g \in L^1(\Omega, \mathcal{F}, \mu)$ and $|f| \le g$ on $\Omega$, then $f \in L^1(\Omega, \mathcal{F}, \mu)$ and $\int_\Omega |f| d\mu \le \int_\Omega g d\mu$.*

*2) If $\mu(\Omega) < \infty$, and if $f$ is $\mathcal{F}$-measurable and bounded on $\Omega$, then $f \in L^1(\Omega, \mathcal{F}, \mu)$.*

**Proof.** 1) is obvious by definition, as both $f^+$ and $f^-$ are dominated by $g$ under assumption. To show 2), suppose $\varphi \in \mathcal{S}^+(\Omega, \mathcal{F})$ and $\varphi \le |f|$. Then $\varphi$ is bounded by $C = \sup_\Omega |f| < \infty$, so that $\varphi \le C 1_\Omega$ on $\Omega$. Hence

$$\int_\Omega \varphi d\mu \le C\mu(\Omega) < \infty.$$

Hence

$$\int_\Omega |f| d\mu = \sup\left\{\int_\Omega \varphi d\mu : \varphi \text{ is non-negative, simple; and } \varphi \le |f|\right\} \le C\mu(E) < \infty.$$

Thus $f \in L^1(\Omega, \mathcal{F}, \mu)$. ■

The following theorem, called Monotone Convergence Theorem (MCT), is one of the most important results in Lebesgue's Theory of Integration.

**Theorem 6.6** *(*MCT, Lebesgue and B. Levi*) Suppose $f_n : \Omega \to [0, \infty]$ are $\mathcal{F}$-measurable and $f_n \uparrow$ (that is $f_{n+1}(x) \ge f_n(x)$ for $n = 1, 2, \cdots$, and $x \in \Omega$). Let $f = \lim_{n\to\infty} f_n$. Then*

$$\int_\Omega f d\mu = \lim_{n\to\infty} \int_\Omega f_n d\mu \ . \tag{6.3}$$

**Proof.** *[The proof is not examinable. We use the proof in W. Rudin: Real and Complex Analysis, Third Edition, page 21]*. Since $f_n \uparrow f$, $f$ is measurable and takes values in $[0, \infty]$. Moreover $f_n \le f_{n+1} \le f$ so that $\int_\Omega f_n d\mu \le \int_\Omega f_{n+1} d\mu \le \int_\Omega f d\mu$. Hence $\lim_{n\to\infty} \int_\Omega f_n d\mu$ exists (but may be $\infty$), and $\lim_{n\to\infty} \int_\Omega f_n d\mu \le \int_\Omega f d\mu$. We next prove the reversed inequality that $\lim_{n\to\infty} \int_\Omega f_n d\mu \ge \int_\Omega f d\mu$.

Suppose $\varphi \in \mathcal{S}^+(\Omega, \mathcal{F})$ and that $\varphi \le f$. Let $\lambda \in (0, 1)$, and let $E_n = \{f_n \ge \lambda\varphi\}$. Since $f_n \uparrow$, $E_n \uparrow$ and $\cup_{n=1}^\infty E_n = \Omega$ . In fact, for any $x \in \Omega$, if $f(x) = 0$ then $x \in E_1$. If $f(x) > 0$, then $\lambda\varphi(x) < f(x)$ (as $\varphi \leq f$ and $\lambda < 1$), since $f_n \uparrow f$, there is $N$ such that $f_n(x) \geq \lambda\varphi(x)$ for all $n \geq N$ (though in general $N$ depends on $x$) so that $x \in E_n$ for $n \geq N$. Thus in any case $x \in \cup_{n=1}^{\infty} E_n$, and therefore $\cup_{n=1}^{\infty} E_n = \Omega$. Since

$$f_n \geq f_n 1_{E_n} \geq \lambda\varphi 1_{E_n}$$

it follows that

$$\int_\Omega f_n d\mu \geq \int_\Omega f_n 1_{E_n} d\mu \geq \int_\Omega \lambda\varphi 1_{E_n} d\mu = \lambda \int_\Omega \varphi 1_{E_n} d\mu. \tag{6.4}$$

Suppose $\varphi = \sum_{i=1}^{k} c_i 1_{A_i}$ where $c_i \geq 0$ and $A_i$ are measurable, then

$$\varphi 1_{E_n} = \sum_{i=1}^{k} c_i 1_{A_i} 1_{E_n} = \sum_{i=1}^{k} c_i 1_{A_i \cap E_n}$$

so that

$$\int_\Omega \varphi 1_{E_n} d\mu = \sum_{i=1}^{k} c_i \mu(A_i \cap E_n).$$

Since $A_i \cap E_n \uparrow A_i$ as $n \to \infty$, thus [Proposition 2.4, item 2)] $\mu(A_i \cap E_n) \to \mu(A_i)$ as $n \to \infty$ for $i = 1, \cdots, k$. Therefore

$$\int_\Omega \varphi 1_{E_n} d\mu \to \sum_{i=1}^{k} c_i \mu(A_i) = \int_\Omega \varphi d\mu \quad \text{as } n \uparrow \infty.$$

Letting $n \to \infty$ in (6.4) we obtain $\lim_{n\to\infty} \int_\Omega f_n d\mu \geq \lambda \int_\Omega \varphi d\mu$ for every $\varphi \in \mathcal{S}^+(\Omega, \mathcal{F})$ such that $\varphi \leq f$ on $\Omega$, which implies that

$$\lim_{n\to\infty} \int_\Omega f_n d\mu \geq \lambda \int_\Omega f d\mu.$$

Since $\lambda \in (0,1)$ arbitrary, we must have $\lim_{n\to\infty} \int_\Omega f_n d\mu \geq \int_\Omega f d\mu$. Therefore $\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega f d\mu$. ■

Let us draw several useful consequences which follow directly from MCT.

**Corollary 6.7** *Suppose $f, g : \Omega \to [0, \infty]$ are measurable.*

*1) There is a sequence of simple non-negative functions $\varphi_n \in \mathcal{S}^+(\Omega, \mathcal{F})$ such that $\varphi_n \uparrow f$ on $\Omega$ and $\int_\Omega f d\mu = \lim_{n\to\infty} \int_\Omega \varphi_n d\mu$.*

*2) We have*

$$\int_\Omega (f+g) d\mu = \int_\Omega f d\mu + \int_\Omega g d\mu$$

*and*

$$\int_\Omega \lambda f d\mu = \lambda \int_\Omega f d\mu$$

*for every number $\lambda \geq 0$.*

**Proof.** 1) follows from Theorem 5.9 and MCT directly. By Theorem 5.9, we may choose $\varphi_n, \psi_n \in \mathcal{S}^+(\Omega, \mathcal{F})$ such that $\varphi_n \uparrow f$ and $\psi_n \uparrow g$. Then $\varphi_n + \psi_n \uparrow f + g$, so that

$$\begin{aligned}\int_\Omega (f+g)d\mu &= \lim_{n\to\infty} \int_\Omega (\varphi_n + \psi_n)\, d\mu = \lim_{n\to\infty} \left( \int_\Omega \varphi_n d\mu + \int_\Omega \psi_n d\mu \right) \\ &= \lim_{n\to\infty} \int_\Omega \varphi_n d\mu + \lim_{n\to\infty} \int_\Omega \psi_n d\mu = \int_\Omega f d\mu + \int_\Omega g d\mu\end{aligned}$$

which proved 2). ■

**Theorem 6.8** *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Then $L^1(\Omega, \mathcal{F}, \mu)$ is a vector space, and $f \to \int_\Omega f d\mu$ is linear on $L^1(\Omega, \mathcal{F}, \mu)$.*

**Proof.** If $f, g \in L^1(\Omega, \mathcal{F}, \mu)$, then $\int_\Omega f^+ d\mu < \infty$ and $\int_\Omega f^- d\mu < \infty$. Since

$$(f+g)^+ \leq f^+ + g^+, \ \ (f+g)^- \leq f^- + g^-$$

so that $\int_\Omega (f+g)^+ d\mu < \infty$ and $\int_\Omega (f+g)^- d\mu < \infty$, which shows that $f + g \in L^1(\Omega, \mathcal{F}, \mu)$. Moreover

$$(f+g)^+ - (f+g)^- = f^+ - f^- + g^+ - g^-$$

so that

$$(f+g)^+ + f^+ + g^+ = (f+g)^- + f^- + g^-,$$

thus

$$\int_\Omega (f+g)^+ d\mu + \int_\Omega f^+ d\mu + \int_\Omega g^+ d\mu = \int_\Omega (f+g)^- d\mu + \int_\Omega f^- d\mu + \int_\Omega g^- d\mu.$$

Rearrange to obtain

$$\int_\Omega (f+g) d\mu = \int_\Omega f d\mu + \int_\Omega g d\mu\ .$$

Suppose $\lambda \geq 0$ is a constant, then $\lambda f = \lambda f^+ - \lambda f^-$, thus if $f \in L^1(\Omega, \mathcal{F}, \mu)$, then both $\lambda f^+, \lambda f^- \in L^1(\Omega, \mathcal{F}, \mu)$, so that by definition $\lambda f$ is integrable

$$\int_\Omega \lambda f d\mu = \int_\Omega \lambda f^+ d\mu - \int_\Omega \lambda f^- d\mu = \lambda \int_\Omega f d\mu.$$

Since $(-f)^+ = f^-$ and $(-f)^- = f^+$, so that $-f$ is integrable by definition, and

$$\int_\Omega (-f) d\mu = -\int_\Omega f d\mu.$$

Therefore $f \to \int_\Omega f d\mu$ is linear, $L^1(\Omega, \mathcal{F}, \mu)$ is a vector space. ■

**Theorem 6.9** *(*MCT for series of non-negative measurable functions, B. Levi*) Let $f_n : \Omega \to [0, \infty]$ be $\mathcal{F}$-measurable. Then*

$$\int_\Omega \sum_{n=1}^\infty f_n d\mu = \sum_{n=1}^\infty \int_\Omega f_n d\mu.$$

*Therefore, $\sum_{n=1}^\infty f_n \in L^1(\Omega, \mathcal{F}, \mu)$ if and only if $\sum_{n=1}^\infty \int_\Omega f_n d\mu < \infty$.*

**Proof.** Apply MCT to the sequence of partial sum $\sum_{i=1}^{n} f_i \uparrow \sum_{n=1}^{\infty} f_n$ to obtain

$$\int_{\Omega} \sum_{n=1}^{\infty} f_n d\mu = \lim_{n\to\infty} \int_{\Omega} \sum_{i=1}^{n} f_i d\mu = \lim_{n\to\infty} \sum_{i=1}^{n} \int_{\Omega} f_i d\mu = \sum_{n=1}^{\infty} \int_{\Omega} f_n d\mu.$$

■

**Theorem 6.10** *Suppose $E_n$ are $\mathcal{F}$-measurable ($n = 1, 2, \cdots$) and are disjoint. Let $E = \cup_{n=1}^{\infty} E_n$. Suppose $f : E \to [0, \infty]$ is measurable. Then $\int_E f d\mu = \sum_{n=1}^{\infty} \int_{E_n} f d\mu$.*

**Proof.** Let $f_n = \sum_{i=1}^{n} f 1_{E_i}$. Then $f_n$ are measurable and $f_n \uparrow f$ on $E$. Therefore, according to MCT,

$$\begin{aligned} \int_E f d\mu &= \lim_{n\to\infty} \int_E f_n d\mu = \lim_{n\to\infty} \int_E \sum_{i=1}^{n} f 1_{E_i} d\mu \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} \int_E f 1_{E_i} d\mu = \lim_{n\to\infty} \sum_{i=1}^{n} \int_{E_i} f d\mu = \sum_{i=1}^{\infty} \int_{E_i} f d\mu. \end{aligned}$$

■

**Corollary 6.11** *Suppose $h : \mathbb{R} \to [0, \infty]$ is Lebesgue measurable. Define $\mu_h(E) = \int_E h dm$ for any $E \in \mathcal{M}_{Leb}$. Then $\mu_h$ is a measure on $(\mathbb{R}, \mathcal{M}_{Leb})$. $d\mu$ is denoted by $hdm$, and $h$ is called the density of the measure $\mu$ with respect to the Lebesgue measure $m$. The measure $\mu_h$ is absolutely continuous with respect to the Lebesgue measure in the sense that if $m(A) = 0$, then $\mu_h(A) = 0$. Suppose $f : \mathbb{R} \to [0, \infty]$ is Lebesgue measurable, then*

$$\int_{\mathbb{R}} f d\mu_h = \int_{\mathbb{R}} f h dm.$$

Suppose $A \subset \Omega$ is a null set, i.e. $\mu(A) = 0$, then any function $f$ on $A$ is $\mathcal{F}$-measurable on the measure space $(A, A \cap \mathcal{F}, \mu)$. If $\varphi \in \mathcal{S}^+(A, \mathcal{F})$ is a simple function on $A$, with a representation $\varphi = \sum_{i=1}^{k} c_i 1_{A_i}$ where $A_i \subset A$, then

$$\int_A \varphi d\mu = \sum_{i=1}^{k} c_i \mu(A_i) = 0$$

which yields that $\int_A f^+ d\mu = \int_A f^- d\mu = \int_A f d\mu = 0$ for every null set $A$. Thus we must have by definition $\int_A f d\mu = 0$ for any null set $A$, for any function $f$ defined on $A$.

If $f, g : \Omega \to [-\infty, \infty]$ be two functions. Suppose $f$ is $\mathcal{F}$-measurable and suppose $f = g$ almost everywhere on $\Omega$. Let $A = \{f \neq g\}$. Then $\mu(A) = 0$. $g$ is $\mathcal{F}$-measurable too. Thus both $A$ and $\Omega \setminus A$ are $\mathcal{F}$-measurable. Now

$$\int_{\Omega} g^+ d\mu = \int_{\Omega\setminus A} g^+ d\mu + \int_A g^+ d\mu = \int_{\Omega\setminus A} g^+ d\mu = \int_{\Omega\setminus A} f^+ d\mu = \int_{\Omega} f^+ d\mu$$

and similarly, $\int_{\Omega} f^- d\mu = \int_{\Omega} g^- d\mu$. Therefore $f \in L^1(\Omega, \mathcal{F}, \mu)$ if and only if $g \in L^1(\Omega, \mathcal{F}, \mu)$, and in this case $\int_{\Omega} f d\mu = \int_{\Omega} g d\mu$. Therefore, the definition of Lebesgue integrals on a measurable set $\Omega$ applies to measurable functions which may be well defined on $\Omega$ only almost surely.

Suppose $E$ and $F$ are two measurable sets, $F \subset E$, and that $E \setminus F$ is null. Suppose $f : E \to [0, \infty]$ is measurable, then

$$\int_E f d\mu = \int_F f d\mu + \int_{E\setminus F} f d\mu = \int_F f d\mu$$

In particular, if $a < b$, then $m([a,b] \setminus (a,b)) = 0$, so that

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

Therefore, we will use $\int_a^b f(x)dx$ to denote the Lebesgue integral $\int_J f dm$ on an interval $J$ with endpoints $a < b$.

**Theorem 6.12** *(*MCT – version for sequences of integrable functions*) Suppose $f_n \in L^1(\Omega, \mathcal{F}, \mu)$ where $n = 1, 2, \cdots$, $f_n \uparrow f$ almost everywhere on $\Omega$ [That is, there is a null set $N \subset \Omega$ such that $f_n(x) \le f_{n+1}(x)$ for all $n$ and $x \in \Omega \setminus N$, and $f(x) = \lim_{n\to\infty} f_n(x)$ for $x \in \Omega \setminus N$], and suppose that the sequence $\left\{\int_\Omega f_n d\mu\right\}$ is bounded above. Then $f \in L^1(\Omega, \mathcal{F}, \mu)$ and $\int_\Omega f d\mu = \lim_{n\to\infty} \int_\Omega f_n d\mu$.*

To prove the theorem, we apply MCT (the version for non-negative functions) to $f_n - f_1$, to obtain

$$\int_\Omega (f - f_1)d\mu = \lim_{n\to\infty} \int_\Omega (f_n - f_1)d\mu = \lim_{n\to\infty} \int_\Omega f_n d\mu - \int_\Omega f_1 d\mu$$

which is finite, so that $f - f_1 \in L^1(\Omega, \mathcal{F}, \mu)$. Hence $f \in L^1(\Omega, \mathcal{F}, \mu)$ and $\int_\Omega f d\mu = \lim_{n\to\infty} \int_\Omega f_n d\mu$.

**Theorem 6.13** *(*Fatou's lemma*) Suppose $f_n : \Omega \to [0, \infty]$ are $\mathcal{F}$-measurable ($n = 1, 2, \cdots$), then*

$$\int_\Omega \liminf_{n\to\infty} f_n d\mu \le \liminf_{n\to\infty} \int_\Omega f_n d\mu. \tag{6.5}$$

*In particular, if $f_n$ are non-negative and measurable, and $f_n \to f$ almost everywhere as $n \to \infty$, then $\int_\Omega f d\mu \le \liminf_{n\to\infty} \int_\Omega f_n d\mu$.*

**Proof.** *[The proof is not examinable]*. Let $g_n(x) = \inf_{i\ge n} f_i(x)$ for $n = 1, 2, \cdots$, so $g_i \le f_n$ for all $i \ge n$, and $g_n \uparrow \liminf_{n\to\infty} f_n$. Apply MCT to $(g_n)$ we have

$$\int_\Omega \liminf_{n\to\infty} f_n d\mu = \lim_{n\to\infty} \int_\Omega g_n d\mu.$$

On the other hand $\int_\Omega g_n d\mu \le \int_\Omega f_i d\mu$ for all $i \ge n$, so that $\int_\Omega g_n d\mu \le \inf_{i\ge n} \int_\Omega f_i d\mu$ and therefore

$$\int_\Omega \liminf_{n\to\infty} f_n d\mu = \lim_{n\to\infty} \int_\Omega g_n d\mu \le \lim_{n\to\infty} \inf_{i\ge n} \int_\Omega f_i d\mu = \liminf_{n\to\infty} \int_\Omega f_n d\mu.$$

■

*Example.* Suppose $f_n$ are integrable on $(\Omega, \mathcal{F}, \mu)$ and bounded in $L^1(\Omega, \mathcal{F}, \mu)$, that is, $\sup_{n\ge 1} \left[\int_\Omega |f_n| d\mu\right] < \infty$, and suppose $f_n \to f$ almost everywhere, then $f$ is integrable. In fact, by Fatou's lemma (applying to the sequence $\{|f_n| : n = 1, 2, \cdots\}$)

$$\int_\Omega |f| d\mu = \int_\Omega \lim_{n\to\infty} |f_n| d\mu \le \liminf_{n\to\infty} \int_\Omega |f_n| d\mu \le \sup_{n\ge 1} \left[\int_\Omega |f_n| d\mu\right] < \infty$$

so that $f \in L^1(\Omega, \mathcal{F}, \mu)$.

**Theorem 6.14** *(*Lebesgue's Dominated Convergence Theorem, DCT*) Let $f_n : \Omega \to [-\infty, \infty]$ be $\mathcal{F}$-measurable, and $f = \lim_{n\to\infty} f_n$ almost everywhere on $\Omega$. Suppose there is an integrable function $g \in L^1(\Omega, \mathcal{F}, \mu)$ such that $|f_n(x)| \le g(x)$ for almost all $x \in \Omega$ for $n = 1, 2, \cdots$. Then*

*1) $f$ and $f_n$ are integrable, and*

*2) $\int_\Omega f d\mu = \lim_{n\to\infty} \int_\Omega f_n d\mu$.*

**Proof.** $f$ is $\mathcal{F}$-measurable, and $|f| \le g$ almost everywhere on $\Omega$, so by comparison, $f_n$, $f$ are integrable. Apply Fatou's lemma to $g - f_n$ to obtain

$$\begin{aligned} \int_\Omega (g-f) \quad d\mu &= \int_\Omega \lim_{n\to\infty} (g - f_n) d\mu \le \liminf \int_\Omega (g - f_n) d\mu \\ &= \int_\Omega g d\mu - \limsup \int_\Omega f_n d\mu \end{aligned}$$

so that

$$\limsup \int_\Omega f_n d\mu \le \int_\Omega f d\mu. \tag{6.6}$$

Apply the inequality above to $-f_n \to -f$ we obtain

$$\limsup \int_\Omega (-f_n)\, d\mu \le \int_\Omega (-f)\, d\mu$$

which is equivalent to that

$$\liminf \int_\Omega f_n d\mu \ge \int_\Omega f d\mu. \tag{6.7}$$

Putting (6.6) and (6.7) together we have

$$\limsup_{n\to\infty} \int_\Omega f_n d\mu \le \int_\Omega f d\mu \le \liminf_{n\to\infty} \int_\Omega f_n d\mu.$$

Since

$$\liminf_{n\to\infty} \int_\Omega f_n d\mu \le \limsup_{n\to\infty} \int_\Omega f_n d\mu$$

so that

$$\liminf_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega f d\mu = \limsup_{n\to\infty} \int_\Omega f_n d\mu$$

hence $\lim \int_\Omega f_n d\mu$ exists and $\lim \int_\Omega f_n d\mu = \int_\Omega f d\mu$, so the proof is complete. ∎

The integrable function $g$ appearing in the previous theorem is called a control function for the sequence $(f_n)$. It is essential to have a control $g$ in DCT. Here is an example. $E = [0, \infty)$ with the Lebesgue measure $m$, $f_n = 1_{[n-1,n)}$. Then $f_n \to 0$ on $E$, $\int_E f_n dm = 1$ for every $n$, but $\lim_{n\to\infty} \int_E f_n dm = 1 \ne \int_E \lim_{n\to\infty} f_n dm$. This example also shows that the condition that $m(E) < \infty$ in the following Corollary is essential.

**Corollary 6.15** *(*Bounded Convergence Theorem, Lebesgue*) Let $(\Omega, \mathcal{F}, \mu)$ be a* finite*, complete measure space, so that $\mu(\Omega) < \infty$. Suppose $f_n$ are measurable, $f_n \to f$ almost everywhere, and $|f_n(x)| \le K$ for all $n$ and for almost all $x \in \Omega$, where $K$ is a constant, then $f_n$ and $f$ are integrable and*

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

We end this part by proving that a Riemann integrable function $f$ on $[a,b]$ must be Lebesgue integrable.

**Theorem 6.16** *Suppose $f : [a,b] \to \mathbb{R}$ is Riemann integrable, then $f$ must be measurable and Lebesgue integrable on $[a,b]$. Moreover, the Lebesgue integral $\int_{(a,b)} f dm$ coincides with the Riemann integral of $f$ over $[a,b]$.*

**Proof.** Recall that $\varphi$ is a step functions if $\varphi = \sum_{i=1}^{k} c_i 1_{J_i}$ where $c_i$ are constants and $J_i$ are finite intervals, which is a simple function on $(\mathbb{R}, \mathcal{M}_{\text{Leb}})$. $I(\varphi)$ denotes the Riemann integral of the step function, that is,

$$I(\varphi) = \sum_{i=1}^{k} c_i |J_i| = \sum_{i=1}^{k} c_i m(J_i) = \int_{\mathbb{R}} \varphi dm$$

according to the definition of Lebesgue measure $m$. Therefore, for step functions, the Lebesgue integrals coincide with their Riemann integrals.

If $f$ is Riemann integrable, then it is bounded. For every $n$ there are two step functions $\varphi_n$ and $\psi_n$ such that

$$\varphi_n \le f 1_{(a,b)} \le \psi_n, \quad I(\varphi_n) \le \int_a^b f(x)dx \le I(\psi_n)$$

and

$$0 \le I(\psi_n) - I(\varphi_n) < \frac{1}{n}.$$

Since the class of step functions is stable under the lattice operations $\wedge$ and $\vee$, that is, if $\varphi_1$ and $\varphi_2$ are step functions, then so are $\varphi_1 \wedge \varphi_2$ and $\varphi_1 \vee \varphi_2$. Therefore we can choose two sequences of step functions $\varphi_n \uparrow$ and $\psi_n \downarrow$ . Let $g = \lim_{n \longrightarrow \infty} \varphi_n$ and $h = \lim_{n \to \infty} \psi_n$. Then both $g$ and $h$ are measurable, and

$$I(\varphi_n - \varphi_1) = I(\varphi_n - \varphi_1) \le \int_a^b f(x)dx - I(\varphi_1)$$

so that, according to MCT, $g - \varphi_1 \in L^1(a,b)$ and therefore $g \in L^1(a,b)$. Similarly $h \in L^1(a,b)$. Moreover

$$I(\psi_n) - I(\varphi_n) < \frac{1}{n}$$

so that, again by MCT

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

Since $h - g \ge 0$ and $\int_{(a,b)} (h-g) dm = 0$, $h = g$ almost every on $(a,b)$. However $g \le f \le h$ so that $f = g = h$ almost surely on $(a,b)$, thus $f$ is Lebesgue measurable. Since $f$ is bounded, and $m((a,b)) = b - a < \infty$, $f$ is (Lebesgue) integrable on $(a,b)$. Moreover

$$\int_{(a,b)} f dm = \int_{(a,b)} h dm = \int_{(a,b)} g dm = \int_a^b f(x)dx,$$

that is the Riemann integral of $f$ on $[a,b]$ coincides with its Lebesgue integral on $[a,b]$ (or equivallently on $(a,b)$). ■

## 6.2 Integrals depending on parameters

The convergence theorems, in particular the Dominated Convergence Theorem, may be applied to a family of measurable functions rather than sequences of measurable functions.

Let $E \subseteq \mathcal{M}_{\text{Leb}}$ be a measurable subset, and $G \subset \mathbb{R}$. Suppose for every $t \in G$, $f_t : E \to [-\infty, \infty]$ is integrable, so that we may form a function $F : G \to \mathbb{R}$ by taking $F(t) = \int_E f_t(x)dx$ for every $t \in G$. We seek for sufficient conditions to ensure that $F$ is continuous on $G$. Let us begin with a simple example.

*Example 1.* Consider $f_t(x) = te^{-t^2x^2}$ for $t \in \mathbb{R}$ and $x \in \mathbb{R}$ which is continuous in $(t, x)$. Then $f_0 = 0$ so that

$$F(0) = \int_{-\infty}^{\infty} f_0(x)\mathrm{d}x = 0$$

while if $t \neq 0$ we have

$$F(t) = \int_{-\infty}^{\infty} te^{-t^2x^2}\mathrm{d}x = \int_{-\infty}^{\infty} e^{-x^2}\mathrm{d}x = \sqrt{2\pi}$$

and therefore $F(t) = \int_{-\infty}^{\infty} te^{-t^2x^2}\mathrm{d}x$ is not continuous at $t = 0$.

**Theorem 6.17** *Let $E$ be measurable, and $J \subset \mathbb{R}$ be an interval. For $t \in J$, $f_t : E \to [-\infty, \infty]$ is measurable. Suppose*

*1) for every $t_0 \in J$, $f_t \to f_{t_0}$ almost everywhere on $E$, and*

*2) there is $g \in L^1(E)$ such that $|f_t| \leq g$ almost everywhere on $E$ for all $t \in J$ (that is, there is a null subset $A \subseteq E$, such that $|f_t(x)| \leq g(x)$ for all $x \in E \setminus A$ and for all $t$).*

*Then $f_t \in L^1(E)$ for every $t$, and $F(t) = \int_E f_t(x)dx$ is continuous on $J$.*

**Proof.** By 2), $f_t \in L^1(E)$ for every $t \in J$, so $F(t) = \int_E f_t(x)dx$ is well defined real function on $J$. To see the continuity at $t_0 \in J$, consider any sequence $t_n \in J$ such that $t_n \to t_0$, apply DCT to $f_{t_n}$ we may deduce that

$$F(t_n) = \int_E f_{t_n}(x)dx \to \int_E f_{t_0}(x)dx = F(t_0),$$

i.e. $F(t_n) \to F(t_0)$ for any $t_n \to t_0$, therefore $F$ is continuous at $t_0$. ∎

It is important to notice that the control function $g$ in 2) is independent of the parameter $t$, as required in the DCT.

We may apply the theorem above to $\frac{f_{t+h}-f_t}{h}$ to obtain the following

**Theorem 6.18** *Let $E$ be a measurable set, and $J \subset \mathbb{R}$ be an interval. For each $t \in J$, $f_t : E \to \mathbb{R}$ is measurable, and the following conditions are satisfied:*

*1) for every $t \in J$, $f_t \in L^1(E)$, and define $F(t) = \int_E f_t(x)dx$ for $t \in J$,*

*2) for every $x \in E$, the partial derivative*

$$\frac{\partial}{\partial t} f_t(x) = \lim_{h \to 0} \frac{f_{t+h}(x) - f_t(x)}{h}$$

*exists for every $t \in J$ (here the limit runs over $h \to 0$ such that $t + h \in J$), and*

*3) there is a control function $g \in L^1(E)$ such that*

$$\left|\frac{\partial}{\partial t} f_t\right| \leq g$$

*almost everywhere on $E$ for all $t \in J$. [Here almost everywhere means that there is a null subset $A \subseteq E$, such that $|\frac{\partial}{\partial t} f(t,x)| \leq g(x)$ for $x \in E \setminus A$ and $t \in J$.]*

*Then $F$ is differentiable on $J$ and*

$$F'(t) = \int_E \frac{\partial}{\partial t} f_t(x) dx \ .$$

**Proof.** Suppose $t \in J$ we want to show $F'(t) = \int_E \frac{\partial}{\partial t} f_t(x)dx$. We may assume that there is $\varepsilon > 0$ such that $[t, t+\varepsilon) \subset J$ (similarly consider the case that $(t-\varepsilon, t] \subset J$). Consider

$$g_h(x) = \frac{f_{t+h}(x) - f_t(x)}{h} \quad \text{for } x \in E \text{ and } h \in (0, \varepsilon).$$

By mean value theorem

$$|g_h(x)| \leq g(x) \quad \text{for almost all } x \text{and for all } h \in (0, \varepsilon).$$

By applying DCT to $g_{h_n}$ where $h_n \in (0, \varepsilon)$ and $h_n \to 0$, to obtain

$$\frac{F(t+h_n) - F(t)}{h_n} = \int_E g_{h_n}(x) dx \to \int_E \frac{\partial}{\partial t} f_t(x) dx$$

so that $F'(t)$ exists and $F'(t) = \int_E \frac{\partial}{\partial t} f_t(x)dx$, which completes the proof. ■

## 6.3 Examples

*Examples 1.* The function $f(x) = \frac{1}{x^p}$ (where $p$ is a constant) is Lebesgue integrable on $(0,1)$ if and only if $p < 1$, and $f$ is Lebesgue integrable on $(1, \infty)$ if and only if $p > 1$.

In fact, $f$ is continuous on $(0,1)$ so it is measurable. It is bounded on $(0,1)$ if $p \leq 0$, so $f$ is Riemann integrable on $[0,1]$, and therefore $f \in L^1(0,1)$. Consider any $p$. For $n = 1, 2, \cdots$, let $E_n = [\frac{1}{n}, 1)$ and $f_n = f 1_{E_n}$. Then $0 \leq f_n \uparrow f$ on $(0,1)$. $f$ is continuous on interval $E_n$ and is bounded, so $f$ is (Riemann) integrable on $E_n$. According to MCT

$$\begin{aligned}
\int_0^1 x^{-p} dx &= \lim_{n\to\infty} \int_{\frac{1}{n}}^1 x^{-p} dx \\
&= \lim_{n\to\infty} \begin{cases} \frac{1}{1-p}(1 - \frac{1}{n^{1-p}}) & \text{if } p \neq 1 \\ \ln n & \text{if } p = 1 \end{cases} = \begin{cases} \frac{1}{1-p} & \text{if } p < 1 \\ \infty & \text{if } p \geq 1 \end{cases}
\end{aligned}$$

Therefore $f \in L^1(0,1)$ if $p < 1$, and $f$ is not integrable on $(0,1)$ if $p \geq 1$.

Similarly, $f(x) = x^{-p}$ on $[1, \infty)$ is continuous so it is measurable. $f$ is non-negative. For each $n$ let $E_n = [1, n]$. $f$ is continuous and bounded on $[1, n]$ thus $f$ is (Riemann) integrable on $[1, n]$.

Clearly $f_n = f1_{E_n} \uparrow f$, so according to MCT

$$\begin{aligned}\int_1^\infty f(x)dx &= \lim_{n\to\infty}\int_1^\infty f_n(x)dx = \lim_{n\to\infty}\int_1^n x^{-p}dx \\ &= \lim_{n\to\infty}\begin{cases}\frac{1}{1-p}(n^{-p+1}-1) & \text{if } p\neq 1 \\ \ln n & \text{if } p=1\end{cases} \\ &= \begin{cases}\frac{1}{p-1} & \text{if } p>1 \\ \infty & \text{if } p\leq 1\end{cases}\end{aligned}$$

so that $x^{-p} \in L^1(1,\infty)$ if $p>1$ and $x^{-p}$ is not integrable on $(1,\infty)$ if $p \leq 1$.

Therefore $x^{-p}$ is not integrable on $(0,\infty)$ for any $p$.

*Example 3.* Let $f(x) = \frac{x^p}{1+x^q}$ where $q \geq 0$. $f$ is integrable on $(1,\infty)$ if and only if $q-p>1$, and $f$ is integrable on $(0,1)$ if and only if $p>-1$. Therefore $f$ is integrable on $(0,\infty)$ if and only if $p>-1$ and $q-p>1$.

*Example 4.* (a) $f(x) = \frac{\sin x}{x}$ is integrable on $(0,1)$, and (b) $f(x) = \frac{\sin x}{x}$ is not integrable on $(0,\infty)$.

By means of contour integral, we have evaluated the (*improper* Riemann) integral

$$\int_0^\infty \frac{\sin x}{x}dx = \lim_{R\to\infty}\int_0^R \frac{\sin x}{x}dx = \frac{\pi}{2}$$

[for example, on page 115, 2.7 Example, in J. B. Conway: *Functions of one complex variable I*]. $\frac{\sin x}{x}$ is continuous and bounded on $(0,R]$ for any $R>0$. Therefore it is (Riemann) integrable on $(0,R]$. The existence of the limit $\lim_{R\to\infty}\int_0^R \frac{\sin x}{x}dx$ doesn't necessarily imply that $f$ is integrable on $(0,\infty)$. In fact, for any $n=1,2,\cdots$, consider $g_n = |f|1_{E_n} \uparrow |f|$ where $E_n = (0,n\pi)$, thus, according to MCT

$$\int_0^\infty |f(x)|dx = \lim_{n\to\infty}\int_0^\infty g_n(x)dx = \lim_{n\to\infty}\int_0^{n\pi}\frac{|\sin x|}{x}dx.$$

Since

$$\begin{aligned}\int_0^{n\pi}\frac{|\sin x|}{x}dx &= \sum_{k=0}^{n-1}\int_{k\pi}^{(k+1)}\frac{|\sin x|}{x}\mathrm{d}x \geq \sum_{k=0}^{n-1}\int_{k\pi}^{(k+1)}\frac{|\sin x|}{(k+1)\pi}\mathrm{d}x \\ &= \sum_{k=0}^{n-1}\frac{1}{(k+1)\pi}\int_0^\pi \sin x\mathrm{d}x = 2\sum_{k=0}^{n-1}\frac{1}{(k+1)\pi} \to \infty\end{aligned}$$

so that $\int_0^\infty |f(x)|dx = \infty$, thus $f$ is not integrable on $(0,\infty)$.

In the following exercises, $E \subset \mathbb{R}$ is a Lebesgue measurable subset, and $m$ is the Lebesgue measure. $L^1(E)$ is the linear space of Lebesgue integrable functions on $E$.

*Example 5.* Use definition to prove carefully the following claims.

a) If $m(E)=0$, then $\int_E \varphi dm = 0$ for any simple function $\varphi$ on $E$.

b) Hence, prove that if $m(E)=0$, then $\int_E f dm = 0$ for any function $f$.

*Proof.* a) First prove that $\int_E \varphi dm = 0$ for any simple measurable function $\varphi$ on $E$. In fact, if $\varphi = \sum_{i=1}^N c_i 1_{A_i}$ where $c_i \geq 0$ and $A_i \subseteq E$. Since $E$ is null, so are $A_i$, thus $m(A_i) = 0$. Therefore

$$\int_E \varphi dm = \sum_{i=1}^N c_i m(A_i) = 0.$$

b) If $m(E) = 0$, then any function $f$ on $E$ is measurable. Since

$$\int_E f^+ dm = \sup\left\{\int_E \varphi dm : \varphi \in S^+(E), \varphi \leq f^+\right\} = 0$$

and by the same reason, $\int_E f^- dm = 0$. Therefore $f \in L^1(E)$, and $\int_E f dm = \int_E f^+ dm - \int_E f^- dm = 0$.

*Example 6.* a) Suppose $f : E \to [0, \infty]$. Prove that $\{f > 0\} = \cup_n^\infty E_n$ where $E_n = \left\{f \geq \frac{1}{n}\right\}$. Hence prove that if in addition $\int_E f dm = 0$, then $m(\left\{f \geq \frac{1}{n}\right\}) = 0$ for any $n = 1, 2, \cdots$, and conclude that $m(\{f > 0\}) = 0$.

b) If $f \in L^1(E)$ then $f$ is finite almost everywhere.

c) If $f$ is measurable, and $\int_E |f| dm = 0$, then $f = 0$ almost everywhere on $E$. If $f, g \in L^1(E)$ and $\int_E |f - g| = 0$, then $f = g$ almost everywhere on $E$.

*Proof.* a) For every $\lambda > 0$, since $f$ is non-negative and measurable, we have

$$m\left[f \geq \lambda\right] = \int_{\{f\geq\lambda\}} dm = \int_{\{f\geq\lambda\}} \frac{f}{\lambda} dm \leq \frac{1}{\lambda}\int_E f dm$$

which is called the Markov inequality. Therefore, if $\int_E f dm = 0$, then for every $\lambda > 0$, $m(\{f \geq \lambda\}) = 0$, so that $m(\{f > 0\}) \leq \sum_{n=1}^\infty m(E_n) = 0$. Therefore $m(\{f > 0\}) = 0$.

b) Suppose $f \in L^1(E)$ and non-negative, then for every $n = 1, 2, \cdots$, $m(\{f \geq n\}) \leq \frac{1}{n}\int_E f dm$. Now $\{f = \infty\} = \cap_{n=1}^\infty \{f \geq n\}$, so that $m(\{f = \infty\}) \leq \frac{1}{n}\int_E f dm \to 0$ as $n \to \infty$, thus $m(\{f = \infty\}) = 0$. For general $f \in L^1(E)$, we apply what proved to $|f|$.

c) follows a) immediately.

*Example 7.* If $f \in L^1(E)$ then

$$\left|\int_E f dm\right| \leq \int_E |f|\, dm.$$

*Hint*: This is because $\int_E f dm = \int_E f^+ dm - \int_E f^- dm$ and $\int_E |f|\, dm = \int_E f^+ dm + \int_E f^- dm$.

*Example 8.* [*Absolute continuity, uniform integrability*] Let $E \subseteq \mathbb{R}$ be a measurable subset.

a) If $f \in L^1(E)$, then $\int_{E_n} f dm \to 0$ as $n \to \infty$, where $E_n = \{x \in E : |f(x)| \geq n\}$.

b) If $f \in L^1(E)$, then the mapping $A \to \int_A f dm$ is *absolutely continuous* with respect to the measure $m$ in the following sense: for any $\varepsilon > 0$ there is $\delta > 0$ such that $\left|\int_A f dm\right| < \varepsilon$ for any measurable subset $A \subseteq E$ such that $m(A) < \delta$.

*Proof.* a) Since $\left|\int_{E_n} f dm\right| \leq \int_{E_n} |f|\, dm$ so without losing generality, we can assume that $f$ is non-negative. Let $f_n = f 1_{E\setminus E_n} = f 1_{\{|f|<n\}}$. The fact that $f$ is finite almost everywhere on $E$

implies that $f_n \uparrow f$ almost everywhere on $E$. Hence, by MCT, $\int_{E \setminus E_n} f dm \to \int_E f dm$ as $n \to \infty$, so that $\int_{E_n} f dm = \int_E f dm - \int_{E\setminus E_n} f dm \to 0$.

b) Since $\left|\int_A f dm\right| \le \int_A |f| \, dm$ so we may assume that $f$ is non-negative in order to prove b). Given $\varepsilon > 0$, choose $N$ such that $\int_{\{f \ge n\}} f dm < \frac{\varepsilon}{2}$ for all $n \ge N$, and use the following (in-)equalities

$$\begin{aligned} \int_A f dm &= \int_{A \cap \{f < N\}} f dm + \int_{A \cap \{f \ge N\}} f dm \\ &\le Nm(A) + \int_{\{f \ge N\}} f dm \le Nm(A) + \frac{\varepsilon}{2}. \end{aligned}$$

Now $\delta = \frac{\varepsilon}{2N}$ will do.

*Example 9.* (*MCT and DCT, almost everywhere versions*). a) Suppose $f_n \in L^1(E)$ and $f_n \uparrow f$ almost everywhere on $E$. If $\int_E f_n dm$ is bounded above, i.e. there is a constant $K$ such that $\int_E f_n \le K$ for all $n$, then $f \in L^1(E)$ and

$$\int_E f = \lim_{n \to \infty} \int_E f_n \ .$$

*Hint.* a) Let $A = \{f_n \uparrow f\}$. Then $A$ is measurable and $E \setminus A$ is a null set. Apply MCT to $f_n - f_1 \uparrow f - f_1$ on $A$ to conclude that $f - f_1 \in L^1(A)$ and therefore $f \in L^1(E)$.

b) Suppose $f_n$, and $f$ are measurable functions on $E$, and $f_n \to f$ almost everywhere on $E$ as $n \to \infty$. Suppose there is a control function $g \in L^1(E)$ such that $|f_n| \le g$ almost everywhere on $E$ for $n = 1, 2, \cdots$. Then $\lim_{n\to\infty} \int_E f_n dm = \int_E f dm$.

c) If $f_n : E \to [-\infty, \infty]$ are measurable, where $E \subseteq \mathbb{R}$ is measurable, $n = 1, 2, \cdots$. Suppose $\sum_{n=1}^\infty \int_E |f_n| < \infty$ then $\sum_{n=1}^\infty f_n$ converges almost everywhere on $E$, $\sum_{n=1}^\infty f_n \in L^1(E)$ and

$$\int_E \sum_{n=1}^\infty f_n dm = \sum_{n=1}^\infty \int_E f_n dm.$$

*Proof.* According to MCT (series version, see Theorem 6.9)

$$\int_E \sum_{n=1}^\infty |f_n| dm = \sum_{n=1}^\infty \int_E |f_n| dm < \infty$$

hence $\sum_{n=1}^\infty |f_n| \in L^1(E)$. In particular, $\sum_{n=1}^\infty |f_n|$ is finite almost everywhere on $E$ [Example 6, b)], that is, $\sum_{n=1}^\infty |f_n|$ converges almost everywhere on $E$. Now consider the sequence $\{S_n\}$ of partial sums, $S_n = \sum_{i=1}^n f_i$. Then $S_n \to S = \sum_{n=1}^\infty f_n$ exists almost everywhere on $E$, and $|S_n| \le \sum_{n=1}^\infty |f_n|$ almost everywhere on $E$ for all $n$, thus we may apply DCT (part b) to $\{S_n\}$ to obtain that

$$\int_E \sum_{n=1}^\infty f_n dm = \lim_{n\to\infty} \int_E \sum_{k=1}^n f_k dm = \lim_{n=\infty} \sum_{k=1}^n \int_E f_k dm = \sum_{n=1}^\infty \int_E f_k dm.$$

*Example 10.* The Gamma function $\Gamma(\alpha)$. Let $f(x) = x^{\alpha-1} e^{-x}$. $f$ is continuous on $(0, \infty)$ so it is measurable. Since $|f(x)| \le x^{\alpha-1}$ on $(0, 1]$, so $f$ is integrable if $\alpha > 0$. While for $x \ge 1$ we have

$$e^{-x} = \frac{1}{\sum_{n=0}^\infty \frac{1}{n!} x^n} \le \frac{n!}{x^n}$$

for any $n$, chose $n$ such that $n - \alpha + 1 \geq 2$, then

$$|f(x)| \leq x^{\alpha-1}\frac{n!}{x^n} \leq \frac{n!}{x^2} \quad \text{for } x \geq 1,$$

hence $f(x)$ is integrable on $[1, \infty)$ for every $\alpha$. Therefore $f$ is integrable on $(0, \infty)$ if $\alpha > 0$. Define

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x \quad \text{for } \alpha > 0.$$

For $n = 1, 2, \cdots$, consider $f_n(x) = x^{\alpha-1}e^{-x}1_{[\frac{1}{n},n]}$. Then $f_n \uparrow f$ so, according to MCT

$$\int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x = \lim_{n\to\infty}\int_{\frac{1}{n}}^n x^{\alpha-1}e^{-x}\mathrm{d}x \text{ [Riemann integral]}$$

On the other hand, using integration by parts we have

$$\begin{aligned}\int_{\frac{1}{n}}^n x^{\alpha-1}e^{-x}\mathrm{d}x &= -\int_{\frac{1}{n}}^n x^{\alpha-1}\mathrm{d}e^{-x} \\ &= -x^{\alpha-1}e^{-x}\Big|_{\frac{1}{n}}^n + (\alpha-1)\int_{\frac{1}{n}}^n x^{(\alpha-1)-1}e^{-x}\mathrm{d}x,\end{aligned}$$

so that, by MCT,

$$\int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x = (\alpha-1)\int_0^\infty x^{(\alpha-1)-1}e^{-x}\mathrm{d}x$$

if $\alpha - 1 > 0$. Hence

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha) \quad \text{for } \alpha > 0.$$

Suppose $\alpha = [\alpha - 1] + r$ where $r \in [1, 0)$, then

$$\Gamma(\alpha) = (\alpha-1)\cdots(\alpha-1-[\alpha-1])\Gamma(r).$$

In particular, if $n$ is positive integer then

$$\begin{aligned}\Gamma(n+1) &= \int_0^\infty x^n e^{-x}\mathrm{d}x = n\Gamma(n) = \cdots = n!\Gamma(1) \\ &= n!\int_0^\infty e^{-x}\mathrm{d}x = n!\ .\end{aligned}$$

The Gamma function $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x$ is continuous on $(0, \infty)$. In fact for every any $\alpha_0 > 0$, choose $J = (a, b)$ such that $0 < a < \alpha_0 < b$, and consider $f_\alpha(x) = x^{\alpha-1}e^{-x}$ for $x \in (0, \infty)$ and $\alpha \in (a, b)$. Let

$$g(x) = x^{a-1}1_{(0,1]} + x^{b-1}e^{-x}1_{(1,\infty)}.$$

Then $g \in L^1(0, \infty)$ and $|f_\alpha(x)| \leq g(x)$ for all $x \in (0, \infty)$ and $\alpha \in (a, b)$. According to Theorem 6.17, $\Gamma(\alpha)$ is continuous at $\alpha_0$. Since $\alpha_0 > 0$ is arbitrary, so $\Gamma$ is continuous on $(0, \infty)$.

*Example 11.* If $\alpha > 0$ then $x^{\alpha-1}e^{-x}$ is integrable on $(0, 1]$. Moreover

$$x^{\alpha-1}e^{-x} = \sum_{n=0}^\infty \frac{(-1)^n}{n!}x^{n+\alpha-1}.$$

Let $h_n(x) = \frac{(-1)^n}{n!} x^{n+\alpha-1}$ which is integrable on $[0,1]$ for any $n = 0, 1, 2, \cdots$. Moreover

$$\int_0^1 |h_n| = \frac{1}{n!}\int_0^1 x^{n+\alpha-1} dx = \frac{1}{n!(n+\alpha)}$$

so that

$$\sum_{n=0}^{\infty}\int_0^1 |h_n| = \sum_{n=0}^{\infty}\frac{1}{n!(n+\alpha)} < \infty.$$

Thus, according to MCT (series version)

$$\int_0^1 x^{\alpha-1}e^{-x}\mathrm{d}x = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\int_0^1 x^{n+\alpha-1}\mathrm{d}x = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!(n+\alpha)}.$$

*Example 12.* Consider function $f(x) = \frac{1}{x}\sin\frac{1}{x}$. For $\varepsilon \in (0,1)$ the function is continuous on $[\varepsilon, 1]$ so it is (Riemann) integrable on $[\varepsilon, 1]$, and

$$\begin{aligned}\int_\varepsilon^1 \frac{1}{x}\sin\frac{1}{x}\mathrm{d}x &= -\int_{\frac{1}{\varepsilon}}^1 \frac{1}{t}\sin t\mathrm{d}t \text{ [substitute } x = \frac{1}{t}]\\ &= \int_1^{\frac{1}{\varepsilon}} \frac{\sin t}{t}\mathrm{d}t \to \int_1^{\infty}\frac{\sin t}{t}\mathrm{d}t\end{aligned}$$

exists as $\varepsilon \downarrow 0$. On the other hand

$$\int_\varepsilon^1 \frac{1}{x}\left|\sin\frac{1}{x}\right|\mathrm{d}x = \int_1^{\frac{1}{\varepsilon}}\frac{|\sin t|}{t}\mathrm{d}t \to \infty$$

as $\varepsilon \downarrow 0$, so according to MCT, $f$ is not integrable on $(0,1]$.

If $f : E \to \mathbb{C}$ is a complex function, then $f \in L^1(E)$ if and only if $\mathrm{Re}f$ and $\mathrm{Im}f$ are integrable on $E$. In this case the Lebesgue integral

$$\int_E f dm = \int_E \mathrm{Re}f dm + i\int_E \mathrm{Im}f dm.$$

It is easy to see that a measurable complex function $f$ on $E$ is integrable if and only if $\int_E |f| < \infty$.

It is easy to see the DCT, hence its corollaries, may apply to complex functions.

*Example 13.* Let $f \in L^1(\mathbb{R})$, and let $f_y(x) = e^{iyx}f(x)$ for $y \in \mathbb{R}$. $f_y$ is measurable and $|f_y| = |f|$, thus $f_y \in L^1(\mathbb{R})$ for every $y$. Therefore, according to Theorem 6.17

$$\tilde{f}(y) = \int_{-\infty}^{\infty} e^{iyx}f(x)\mathrm{d}x$$

is continuous on $\mathbb{R}$.

Let $g_y(x) = \frac{e^{iyx}-1}{ix}f(x)$. Then

$$\begin{aligned}|g_y(x)| &= \frac{|e^{iyx}-1|}{|x|}f(x) = \frac{|2\sin(\frac{yx}{2})|}{|x|}|f(x)|\\ &\leq |y||f(x)| \qquad \text{for} \quad x \neq 0,\end{aligned}$$

so that $g_y \in L^1(\mathbb{R})$. Moreover $\frac{\partial}{\partial y} g_y(x) = e^{iyx} f(x)$ and

$$\left| \frac{\partial}{\partial y} g_y(x) \right| = |f(x)| \quad \forall y \text{ and } x.$$

By Theorem 6.18 we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}y} \int_{-\infty}^{\infty} g_y(x) \mathrm{d}x &= \int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}y} \frac{e^{iyx} - 1}{ix} f(x) \mathrm{d}x \\ &= \int_{-\infty}^{\infty} e^{iyx} f(x) \mathrm{d}x = \tilde{f}(y). \end{aligned}$$

That is

$$\tilde{f}(y) = \frac{\mathrm{d}}{\mathrm{d}y} \int_{-\infty}^{\infty} \frac{e^{iyx} - 1}{ix} f(x) \mathrm{d}x.$$

# 7 Fubini's theorem

The Lebesgue measure on $\mathbb{R}^d$, which is the extensions of $d$-dimensional volumes (areas if $d = 2$, volumes if $d \geq 3$) to a large class of measurable subsets, can be constructed in a similar way as we did for $\mathbb{R}$. Another approach is to define the Lebesgue measure $m$ on $\mathbb{R}^2$ as the *product measure*, and repeat the definition of product measures to define the Lebesgue measure on $\mathbb{R}^d$ for general $d$.

## 7.1 Measures on product spaces

Let us explain the idea how to construct the Lebesgue measure on $\mathbb{R}^2$ as the product measure of the Lebesgue measure on $\mathbb{R}$. The same idea actually applies to the general setting of $\sigma$-finite measure spaces only with few modifications. In particular, what we are going to outline applies to Lebesgue sub-spaces.

Let us take two copies of the real line $\mathbb{R}$, but label them differently, so let $X = \mathbb{R}$ and $Y = \mathbb{R}$, use $\mathcal{F}$ to denote the $\sigma$-algebra $\mathcal{M}_{\mathrm{Leb}}$ on $X$ and $\mathcal{G}$ the $\sigma$-algebra $\mathcal{M}_{\mathrm{Leb}}$ on $Y$. Let us use $m_1$ to denote the Lebesgue measure on $(X, \mathcal{F})$ and $m_2$ the Lebesgue measure on $(Y, \mathcal{G})$, respectively. The theory we are going to develop relies only on the fact that $(X, \mathcal{F}, m_1)$ and $(Y, \mathcal{G}, m_2)$ are two ($\sigma$-finite) measure spaces. This is the reason why we label the same Lebesgue measure one as $m_1$ and the other as $m_2$.

The space of all ordered pairs $(x, y)$ where $x \in X$ and $y \in Y$ is called the (Cartesian) product of $X$ and $Y$, denoted by $X \times Y$. For example, $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, and $[0,1] \times [0,1] = [0,1]^2$ is the unit square in $\mathbb{R}^2$.

Suppose $A \in \mathcal{F}$ and $B \in \mathcal{G}$, then $A \times B \subset X \times Y$ is called a measurable rectangle type set, and its measure is naturally defined to be

$$m(A \times B) = m_1(A) m_2(B) \tag{7.1}$$

here we use the convention that $0 \cdot \infty = \infty \cdot 0 = 0$. There are however different approaches to define the product measure $m$ such that (7.1) holds for every measurable rectangle type set. One way which is natural but is not the simplest and most direct way is to define the outer measure

$$m^*(E) = \inf\left\{\sum_{i=1}^{\infty} m_1(A_i)m_2(B_i) : \ A_i \in \mathcal{F}, B_i \in \mathcal{G} \text{ s.t. } \bigcup_{i=1}^{\infty} A_i \times B_i \supseteq E\right\}$$

for any subset $E \subset X \times Y$. Then $m^*$ is a outer measure on the product space $X \times Y$:

1) $m^*(\emptyset) = 0$ and $m^*(E) \geq 0$ if any $E \subseteq X \times Y$;
2) $m^*(E) \leq m^*(F)$ if $E \subseteq F \subseteq X \times Y$;
3) $m^*$ is countably sub-additive, that is,

$$m^*\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} m^*(E_i)$$

for any $E_i \subseteq X \times Y$ where $i = 1, 2, \cdots$.

A subset $E \subset X \times Y$ is $m^*$-measurable if $E$ satisfies Caratheodory's condition:

$$m^*(F) = m^*(F \cap E) + m^*(F \cap E^c)$$

for every $F \subset X \times Y$. Then, the collection of all $m^*$-measurable subsets of the product space $X \times Y$ is a $\sigma$-algebra on $X \times Y$, denoted by $\mathcal{M}_{\text{Leb}}(X \times Y)$ or by $\mathcal{M}_{\text{Leb}}$ if no confusion may arise. The outer measure $m^*$ restricted on $\mathcal{M}_{\text{Leb}}(X \times Y)$ is a measure, which is denoted by $m$.

One can show that $A \times B \in \mathcal{M}_{\text{Leb}}(X \times Y)$ if $A \in \mathcal{F}$ and $B \in \mathcal{G}$, so that

$$\mathcal{F} \otimes \mathcal{G} = \sigma\left\{A \times B : \text{ where } A \in \mathcal{F} \text{ and } B \in \mathcal{G}\right\} \subseteq \mathcal{M}_{\text{Leb}}(X \times Y).$$

$m^*$ restricted on $\mathcal{F} \otimes \mathcal{G} \subseteq \mathcal{M}_{\text{Leb}}(X \times Y)$ is of course a measure, thus again denoted by $m$, and therefore $(X \times Y, \mathcal{F} \otimes \mathcal{G}, m)$ is a measure space, called the *product measure space* of $(X, \mathcal{F}, m_1)$ and $(Y, \mathcal{G}, m_2)$. In general $\mathcal{F} \otimes \mathcal{G}$ is strictly smaller than $\mathcal{M}_{\text{Leb}}(X \times Y)$.

The measure space $(\mathbb{R}^2, \mathcal{M}_{\text{Leb}}(\mathbb{R}^2), m)$ is called the (two-dimensional) Lebesgue measure space.

If $E \subset \mathbb{R}^2$ is $m^*$-measurable, that is, $E \in \mathcal{M}_{\text{Leb}}$, then

$$\mathcal{M}_{\text{Leb}}(E) = \left\{E \cap F : F \in \mathcal{M}_{\text{Leb}}(\mathbb{R}^2)\right\}$$

is a $\sigma$-algebra on $E$, and $\mathcal{M}_{\text{Leb}}(E) \subset \mathcal{M}_{\text{Leb}}(\mathbb{R}^2)$. The restriction of the Lebesgue measure $m$ on $\mathcal{M}_{\text{Leb}}(E)$ is a measure on $(E, \mathcal{M}_{\text{Leb}}(E))$. $(E, \mathcal{M}_{\text{Leb}}(E), m)$ is a two dimensional Lebesgue (sub)-space.

The Lebesgue's theory of integration is applicable to the measure space $(E, \mathcal{M}_{\text{Leb}}(E), m)$ where $E \subset \mathbb{R}^2$ is measurable. The integral of a non-negative measurable function / integrable function $f$ on $E \in \mathcal{M}_{\text{Leb}}$ is denoted by $\int_E f(x,y)\mathrm{d}x\mathrm{d}y$ or by $\int_E f dm$ for simplicity. If $f$ is integrable on $E$, then we say $f \in L^1(E)$, and $\int_E f(x,y)\mathrm{d}x\mathrm{d}y$ is called the Lebesgue integral of $f$ on $E$, also called a *double integral*.

Suppose $E = [a,b] \times [c,d]$ is a closed rectangle in $\mathbb{R}^2$, and $f : E \to \mathbb{R}$ is bounded. Suppose the double integral in the sense of Riemann

$$\iint_E f(x,y)\mathrm{d}x\mathrm{d}y \tag{7.2}$$

exists, then $f \in L^1(E)$ and the Lebesgue integral $\int_E f dm$ coincides with the Riemann double integral (7.2).

## 7.2 Fubini's theorem and Tonelli's theorem

Let $X, Y \in \mathcal{M}_{\text{Leb}}(\mathbb{R})$. Then $X \times Y \in \mathcal{M}_{\text{Leb}}(\mathbb{R}^2)$, and $(X \times Y, \mathcal{M}_{\text{Leb}}(X \times Y), m)$ be a (two dimensional) Lebesgue space. We have the following simple observation.

**Lemma 7.1** *Suppose $\varphi$ is a non-negative simple function on $X \times Y$ of the following form:*

$$\varphi = \sum_{i=1}^{k} c_i 1_{A_i \times B_i} \quad , \quad A_i \in \mathcal{M}_{Leb}(X), B_i \in \mathcal{M}_{Leb}(Y), \;\; c_i \geq 0,$$

*then*

*1) for any $y \in Y$, $\varphi_y \in \mathcal{S}^+(X)$ where $\varphi_y(x) = \varphi(x, y)$, and $F \in \mathcal{S}^+(Y)$ where $F(y) = \int_X \varphi_y dm_1$, and*

*2) we have*

$$\int_Y F dm_2 = \int_{X \times Y} \varphi dm$$

*that is*

$$\int_Y \left( \int_X \varphi(x, y)\, dx \right) dy = \int_X \left( \int_Y \varphi(x, y)\, dy \right) dx = \int_{X \times Y} \varphi(x, y)\, dx\, dy.$$

**Proof.** According to the definition of the Lebesgue measure on $X \times Y$, we have

$$\int_{X \times Y} \varphi(x, y) \mathrm{d}x \mathrm{d}y = \sum_{i=1}^{k} c_i m(A_i \times B_i) = \sum_{i=1}^{k} c_i m_1(A_i) m_2(B_i).$$

On the other hand, for each $y$,

$$\varphi_y = \varphi(\cdot, y) = \sum_{i=1}^{k} c_i 1_{B_i}(y) 1_{A_i}$$

which is a non-negative measurable simple function on $X$, and

$$\begin{aligned} F(y) &= \int_X \varphi_y(x) dx = \int_X \varphi(x, y) \mathrm{d}x = \sum_{i=1}^{k} c_i 1_{B_i}(y) m_1(A_i) \\ &= \sum_{i=1}^{k} c_i m_1(A_i) 1_{B_i}(y) \end{aligned}$$

which is a non-negative simple function on $Y$, and its integral

$$\begin{aligned} \int_Y F dm_2 &= \int_Y \left( \int_X \varphi(x, y) \mathrm{d}x \right) \mathrm{d}y = \sum_{i=1}^{k} c_i m(A_i) m(B_i) \\ &= \int_{X \times Y} \varphi(x, y) \mathrm{d}x \mathrm{d}y. \end{aligned}$$

■

**Proposition 7.2** *Let $f : X \times Y \to [0, \infty)$ be measurable. Then*
*1) $x \to f(x, y)$ is measurable on $X$ for almost all $y \in Y$,*
*2) $y \to \int_X f(x, y)dx$ is (define for almost all $y$) non-negative and measurable on $Y$, and*
*3) we have the equality*

$$\int_{X\times Y} f(x,y)\,dx\,dy = \int_Y \left( \int_X f(x,y)\,dx \right) dy = \int_X \left( \int_Y f(x,y)\,dy \right) dx.$$

The proof is a careful applications of MCT together with the definition of the outer measure $m^*$. We omit the detail of its proof.

**Theorem 7.3** *(Fubini's theorem) Let $X, Y \in \mathcal{M}_{Leb}(\mathbb{R})$ [so $X \times Y \in \mathcal{M}_{Leb}(\mathbb{R}^2)$] and $f \in L^1(X \times Y)$. Then*
*1) for almost all $y \in Y$, $f_y \in L^1(X)$, where $f_y(x) = f(y, x)$ for $x \in X$, so $F(y) = \int_X f_y(x)dx$ is well defined for almost all $y \in Y$, and*
*2) $F$ defined in 1) is integrable on $Y$ (so in particular, $F$ is Lebesgue measurable), and*

$$\int_Y F(y)dy = \int_{X\times Y} f(x,y)dxdy.$$

*Therefore*

$$\int_Y \left( \int_X f(x,y)\,dx \right) dy = \int_X \left( \int_Y f(x,y)\,dy \right) dx = \int_{X\times Y} f(x,y)\,dx\,dy.$$

**Proof.** Apply Proposition 7.2 to $f^+$ and $f^-$. ■

**Theorem 7.4** *(Tonelli's theorem) Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is measurable, and suppose either of the repeated integrals exists and is* finite, *i.e.*

$$\int_{\mathbb{R}} \left( \int_{\mathbb{R}} |f(x,y)|\,dx \right) dy < \infty, \quad \textit{or / and} \int_{\mathbb{R}} \left( \int_{\mathbb{R}} |f(x,y)|\,dy \right) dx < \infty.$$

*Then $f \in L^1(\mathbb{R}^2)$, so that Fubini's theorem is applicable to both $f$ and $|f|$.*

**Proof.** According to Proposition 7.2, under the assumptions, $\int_{\mathbb{R}^2} |f|dm < \infty$, so that $f \in L^1(\mathbb{R}^2)$. ■

## 7.3 Examples

Let us give several examples to demonstrate how to use Fubini's Theorem and Tonelli's Theorem. I took some examples from Batty's notes written for 2013H.

*Example 1.* Let $E = [-1, 1] \times [-1, 1]$ and $f(x, y) = \frac{xy}{(x^2+y^2)^2}$ if $(x, y) \neq (0, 0)$ and $f(0, 0) = 0$. $f$ is measurable on $E$, and for any $y \neq 0$, $x \to \frac{xy}{(x^2+y^2)^2}$ is continuous on $[-1, 1]$ so is integrable in $[-1, 1]$. Moreover

$$\int_{-1}^{1} \frac{xy}{(x^2+y^2)^2} \mathrm{d}x = 0 \quad \text{for any } y \neq 0$$

and therefore

$$\int_{-1}^{1}\left(\int_{-1}^{1}\frac{xy}{(x^2+y^2)^2}\mathrm{d}x\right)\mathrm{d}y=0.$$

Similarly

$$\int_{-1}^{1}\left(\int_{-1}^{1}\frac{xy}{(x^2+y^2)^2}\mathrm{d}y\right)\mathrm{d}x=0.$$

Two repeated integrals exist and coincide, but $f$ is not integrable. In fact, $f$ is not integrable on $(0,1]\times(0,1]\subset E$, since

$$\int_{0}^{1}\frac{xy}{(x^2+y^2)^2}\mathrm{d}y=\frac{1}{2x}-\frac{x}{2(x^2+1)}$$

which is not integrable on $(0,1]$, so by Fubini's theorem, $f$ can not be integrable on $(0,1]^2$, neither on $[-1,1]^2$.

*Example 2.* The repeated integral

$$\int_{0}^{1}\left(\int_{0}^{x}\sqrt{\frac{1-y}{x-y}}\mathrm{d}y\right)\mathrm{d}x$$

may be written as

$$I=\int_{0}^{1}\left(\int_{0}^{1}\sqrt{\frac{1-y}{x-y}}1_{\{y<x\}}\mathrm{d}y\right)\mathrm{d}x$$

so consider $f(x,y)=\sqrt{\frac{1-y}{x-y}}1_{\{y<x\}}$ on $E=[0,1]\times[0,1]$. The function is non-negative and measurable on $E$, so according to Proposition 7.2

$$\int_{0}^{1}\left(\int_{0}^{1}f(x,y)\mathrm{d}y\right)\mathrm{d}x=\int_{0}^{1}\left(\int_{0}^{1}f(x,y)\mathrm{d}x\right)\mathrm{d}y=\int_{E}f.$$

On the other hand

$$\begin{aligned}\int_{0}^{1}\left(\int_{0}^{1}f(x,y)\mathrm{d}x\right)\mathrm{d}y &= \int_{0}^{1}\left(\int_{0}^{1}\sqrt{\frac{1-y}{x-y}}1_{\{y<x\}}\mathrm{d}x\right)\mathrm{d}y\\ &= \int_{0}^{1}\left(\int_{y}^{1}\sqrt{\frac{1-y}{x-y}}\mathrm{d}x\right)\mathrm{d}y\\ &= \int_{0}^{1}\sqrt{1-y}2\sqrt{1-y}\mathrm{d}y=1\end{aligned}$$

so $f\in L^1(E)$ and the repeated integral $I=1$.

*Example 3.* Consider $f(x,y)=ye^{-y^2(1+x^2)}$ on $E=(0,\infty)\times(0,\infty)$ the first quadrant. $f$ is continuous, and non-negative on $E$. Thus $f$ is measurable. The repeated integral

$$\int_{0}^{\infty}\left(\int_{0}^{\infty}ye^{-y^2(1+x^2)}\mathrm{d}y\right)\mathrm{d}x$$

is easy to work out: for $x > 0$ we have

$$\begin{aligned}\int_0^\infty ye^{-y^2(1+x^2)}\mathrm{d}y &= \frac{1}{1+x^2}\int_0^\infty ze^{-z^2}\mathrm{d}z \\ \text{[Making change of variable } \sqrt{1+x^2}y &= z\ ] \\ &= \frac{1}{2}\frac{1}{1+x^2}\end{aligned}$$

so that

$$\int_0^\infty\left(\int_0^\infty ye^{-y^2(1+x^2)}\mathrm{d}y\right)\mathrm{d}x = \int_0^\infty \frac{1}{2}\frac{1}{1+x^2}\mathrm{d}x = \frac{\pi}{4}.$$

Thus $f$ is integrable on $(0,\infty)\times(0,\infty)$. By Fubini's theorem, the other related integral of $f$ on $E$ is $\frac{\pi}{4}$ as well. Thus

$$\begin{aligned}\frac{\pi}{4} &= \int_0^\infty\left(\int_0^\infty ye^{-y^2(1+x^2)}\mathrm{d}x\right)\mathrm{d}y = \int_0^\infty\left(e^{-y^2}\int_0^\infty e^{-y^2x^2}y\mathrm{d}x\right)\mathrm{d}y \\ &= \int_0^\infty\left(e^{-y^2}\int_0^\infty e^{-x^2}\mathrm{d}x\right)\mathrm{d}y = \left(\int_0^\infty e^{-x^2}\mathrm{d}x\right)^2\end{aligned}$$

and therefore

$$\int_0^\infty e^{-x^2}\mathrm{d}x = \frac{\sqrt{\pi}}{2}.$$

Another way to calculate the integral $I = \int_0^\infty e^{-x^2}\mathrm{d}x$ can be described as the following. Consider $I^2$ and write it as

$$\begin{aligned}I^2 &= \int_0^\infty e^{-y^2}\mathrm{d}y\int_0^\infty e^{-x^2}\mathrm{d}x = \int_0^\infty\left(e^{-y^2}\int_0^\infty e^{-x^2}\mathrm{d}x\right)\mathrm{d}y \\ &= \int_0^\infty\left(\int_0^\infty e^{-(x^2+y^2)}\mathrm{d}x\right)\mathrm{d}y.\end{aligned}$$

According to Proposition 7.2 and MCT we have

$$\int_0^\infty\left(\int_0^\infty e^{-(x^2+y^2)}\mathrm{d}x\right)\mathrm{d}y = \int_E e^{-(x^2+y^2)}dxdy = \lim_{n\to\infty}\int_E e^{-(x^2+y^2)}1_{\{x^2+y^2\le n^2\}}dxdy.$$

On the other hand, the integral $\int_E e^{-(x^2+y^2)}1_{\{x^2+y^2\le n^2\}}$ coincides with the Riemann integral which can be calculated in terms of polar coordinates, so that

$$\int_E e^{-(x^2+y^2)}1_{\{x^2+y^2\le n^2\}}dxdy = \iint_{\{(r,\theta):0<r\le n, 0\le\theta\le\frac{\pi}{2}\}} e^{-r^2}r\mathrm{d}r\mathrm{d}\theta = \int_0^{\frac{\pi}{2}}\left(\int_0^n e^{-r^2}r\mathrm{d}r\right)\mathrm{d}\theta$$

where the last equality follows Fubini's theorem. So that

$$\int_E e^{-(x^2+y^2)}1_{\{x^2+y^2\le n^2\}}dxdy = \frac{\pi}{4}(1-e^{-n^2}) \to \frac{\pi}{4}$$

which again gives $I = \sqrt{\pi}/2$.

## 7.4 Change of variables

We have made changes of variables in order to evaluate Lebesgue integrals which take the same form as for Riemann integrals. Let us write down a precise statement.

**Theorem 7.5** *Suppose $\varphi : (a,b) \to (c,d)$ is one to one, onto, and differentiable, $\varphi' \geq 0$. If $f : (c,d) \to [-\infty, \infty]$ is measurable, then $f \in L^1(c,d)$ if and only if $\varphi' f \circ \varphi \in L^1(a,b)$. In this case*

$$\int_c^d f(x)dx = \int_a^b f(\varphi(t))\varphi'(t)dx \ .$$

**Proof.** *[The proof is not examinable]*. By construction of Lebesgue integrals, we only need to consider the case that $f$ is non-negative and Lebesgue measurable. However, any non-negative and measurable function is the limit of an increasing sequence of non-negative simple measurable functions, by MCT and the linearity of the Lebesgue integration, we only need to show the statement is valid for $f = 1_E$ where $E \subset (c,d)$ is a measurable subset. In this case $\int_c^d 1_E dm = m(E) < \infty$. By definition, for every $\varepsilon > 0$ there is a sequence of intervals $(a_n, b_n)$ such that $\cup_{n=1}^\infty (a_n, b_n) \supset E$ and

$$\sum_{n=1}^\infty (b_n - a_n) - \varepsilon \leq m(E) \leq \sum_{n=1}^\infty (b_n - a_n). \tag{7.3}$$

On the other hand,

$$1_E(\varphi(t))\varphi'(t) \leq \sum_{n=1}^\infty 1_{(a_n,b_n)}(\varphi(t))\,\varphi'(t)$$

on $(a,b)$, so that

$$\begin{aligned}\int_a^b 1_E(\varphi(t))\varphi'(t)dt &\leq \sum_{n=1}^\infty \int_a^b 1_{(a_n,b_n)}(\varphi(t))\,\varphi'(t) \\ &\leq \sum_{n=1}^\infty (b_n - a_n)\end{aligned}$$

which follows that, together with (7.3)

$$\int_a^b 1_E(\varphi(t))\varphi'(t)dt \leq m(E) + \varepsilon = \int_c^d 1_E(x)dx + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we have

$$\int_a^b 1_E(\varphi(t))\varphi'(t)dt \leq \int_c^d 1_E(x)dx$$

so that

$$\int_a^b f(\varphi(t))\varphi'(t)dt \leq \int_c^d f(x)dx$$

for non-negative measurable function $f$. Applying the inequality to the inverse of $\varphi$ we also have

$$\int_c^d f(x)dx \leq \int_a^b f(\varphi(t))\varphi'(t)dt$$

so that we must have

$$\int_c^d f(x)dx = \int_a^b f(\varphi(t))\varphi'(t)dt$$

which completes the proof. ■

*Example* 1. Consider the Fourier transform of $e^{-x^2}$. Note that

$$e^{-x^2} \leq \frac{1}{1+x^2}$$

so that $g(x) = e^{-x^2}$ is integrable on $\mathbb{R}$. For $s \in \mathbb{R}$ consider $f(x) = e^{-isx}g(x)$. Then $f$ is measurable and $|f| \leq g$, thus $f$ is integrable. Moreover

$$f(x) = e^{-isx}e^{-x^2} = \sum_{n=0}^{\infty} \frac{(-is)^n}{n!} x^n e^{-x^2}.$$

Now $x^n e^{-x^2}$ is integrable, so that if $n = 2m+1$ is odd then, by DCT we have

$$\begin{aligned}
\int_{-\infty}^{\infty} x^{2m+1}e^{-x^2}\mathrm{d}x &= \lim_{N\to\infty} \int x^{2m+1}e^{-x^2}1_{[-N,N]}\mathrm{d}x \\
&= \lim_{N\to\infty} \int_{-N}^{N} x^{2m+1}e^{-x^2}\mathrm{d}x = 0.
\end{aligned}$$

If $n = 2m$ is even,

$$\int_{-\infty}^{\infty} x^{2m}e^{-x^2}\mathrm{d}x = 2\int_0^{\infty} x^{2m}e^{-x^2}\mathrm{d}x = \int_0^{\infty} z^{m-\frac{1}{2}}e^{-z}\mathrm{d}z$$

[we have made change of variable $z = x^2$ to obtain the last equality], so that

$$\begin{aligned}
\int_{-\infty}^{\infty} x^{2m}e^{-x^2}\mathrm{d}x &= \Gamma(m+\frac{1}{2}) = \int_0^{\infty} z^{m-\frac{1}{2}}e^{-z}\mathrm{d}z \\
&= \left(m-\frac{1}{2}\right)\left(m-1-\frac{1}{2}\right)\cdots\frac{1}{2}\Gamma\left(\frac{1}{2}\right) \\
&= \frac{(2m-1)!!}{2^m}\Gamma\left(\frac{1}{2}\right).
\end{aligned}$$

Hence

$$\begin{aligned}
\sum_{n=0}^{\infty}\int_{-\infty}^{\infty}\left|\frac{(-is)^n}{n!}x^n e^{-x^2}\right| &= \sum_{m=0}^{\infty}\frac{s^{2m}}{(2m)!}\frac{(2m-1)!!}{2^m}\Gamma\left(\frac{1}{2}\right) \\
&= \sum_{m=0}^{\infty}\frac{s^{2m}}{m!}\frac{1}{4^m}\Gamma\left(\frac{1}{2}\right) = e^{\frac{s^2}{4}}\Gamma\left(\frac{1}{2}\right) \\
&< \infty,
\end{aligned}$$

and therefore, according to MCT (series version, Example 9 Part c) in Section 6.4 Examples)

$$\begin{aligned}\tilde{f}(s) &= \int_{-\infty}^{\infty} e^{-isx}e^{-x^2}\mathrm{d}x = \sum_{n=0}^{\infty}\frac{(-is)^n}{n!}\int_{-\infty}^{\infty} x^n e^{-x^2}\mathrm{d}x \\ &= \sum_{m=0}^{\infty}\frac{(-is)^{2m}}{(2m)!}\frac{(2m-1)!!}{2^m}\Gamma\left(\frac{1}{2}\right) = e^{-\frac{s^2}{4}}\Gamma\left(\frac{1}{2}\right).\end{aligned}$$

On the other hand

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} z^{-\frac{1}{2}}e^{-z}\mathrm{d}z = \int_0^{\infty} 2e^{-x^2}\mathrm{d}z = \sqrt{\pi}$$

so that the Fourier transform of $e^{-x^2}$ is $\sqrt{\pi}e^{-\frac{s^2}{4}}$.

**Theorem 7.6** *Let $G$ be an open subset of $\mathbb{R}^2$, and*

$$T : G \to E = T(G) \subset \mathbb{R}^2$$

*be one-to-one, and differentiable. Suppose $f : E \to \mathbb{R}$ is measurable. Then $f$ is integrable if and only if $f \circ T|\det J_T|$ is integrable on $G$, where $\det J_T$ is the Jacobian of $T$. In that case*

$$\int_E f(x,y)d(x,y) = \int_G f \circ T(u,v)|\det J_T(u,v)|dudv$$

Recall that if we write $T(u,v) = (x(u,v), y(u,v))$ then

$$J_T = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}.$$

For example, if we use polar coordinate $T(r,\theta) = (x,y)$, where $x = r\cos\theta$, $y = r\sin\theta$ then $|\det J_T| = r$. Suppose that $E = T(G)$ and $f$ is measurable. Then $f$ is integrable on $E$ if and only if

$$f(r\cos\theta, r\sin\theta)r$$

is integrable on $G$.

# 8 $L^p$-space and convergences

Let $(E, \mathcal{F}, \mu)$ be a *complete* measure space. For example $E$ is a measurable subset of $\mathbb{R}$ or $\mathbb{R}^2$, $\mathcal{F} = \mathcal{M}_{\mathrm{Leb}}(E)$ and $\mu = m$ is the Lebesgue measure on $E$. Another example is the discrete model, where $E = \mathbb{N}$, $\mathbb{Z}_+ = \{0\} \cup \mathbb{N}$ or $\mathbb{Z}$, or $E$ can be any discrete set, $\mathcal{F}$ is the $\sigma$-algebra $\mathcal{P}(E)$ of all subsets of $E$, and $\mu(\{n\}) = p_n$ where $p_n \geq 0$ for all $n \in E$. If in addition $\sum_{n\in E} p_n = 1$, then $\mu$ is a probability measure.

## 8.1 $L^p$-spaces

Suppose $f : E \to [-\infty, \infty]$ is measurable, and suppose $p$ is a positive constant, then $|f|^p$ is measurable too. If $\int_E |f|^p d\mu < \infty$ then $f$ is called $p$-th integrable in $E$. The collection of all such functions on $E$ is denoted by $L^p(E)$. $L^p(E)$ is a vector space (over the real field). More precisely, $L^p(E)$ is the space of equivalence classes under the equivalence $\sim$, where we say $f \sim g$, if $\mu[f \neq g] = 0$, i.e. $f = g$ almost everywhere on $E$. For simplicity, when dealing with $L^p$-spaces, we prefer to identify the equivalence $[f]$ containing $f$ with any $g$ in the equivalence $[f]$. That is, we view two functions $f$ and $g$ as the same element in $L^p(E)$ if $f = g$ almost everywhere (with respect to the measure $\mu$) on $E$.

A complex measurable function $f$ on $E$ is $p$-th integrable, if both Re$f$ and Im$f$ belong to $L^p(E)$, which is again equivalent to that $\int_E |f|^p d\mu < \infty$. The space of all $p$-th integrable measurable complex functions is denoted again by $L^p(E)$ (sorry we use the same notation as for real functions). $L^p(E)$ is then a vector space over the complex field. In fact it is obvious that if $f \in L^p(E)$ then $\alpha f \in L^p(E)$ for any constant $\alpha$. Since

$$|f+g|^p \leq 2^p(|f|^p + |g|^p)$$

so $f + g \in L^p(E)$ if $f, g \in L^p(E)$. Thus $L^p(E)$ is a vector space.

For the case that $p = 1$, define

$$d_1(f,g) = \int_E |f-g| d\mu$$

which is called the $L^1$distance between $f$ and $g$. Since

$$\int_E |f+g| d\mu \leq \int_E |f| d\mu + \int_E |g| d\mu$$

so that the *triangle inequality* holds

$$d_1(f,g) \leq d_1(f,h) + d_1(h,g).$$

Moreover

$$d_1(f,g) = \int_E |f-g| d\mu = 0$$

if and only if $f = g$ almost everywhere on $E$. Therefore, if we identify the functions which are equal almost everywhere on $E$ [i.e. we identify an integrable function $f$ as its equivalent class: if $f = g$ almost everywhere on $E$, then $f$ and $g$ are considered as the same function on $E$], then $L^1(E)$ is a metric space.

We wish to define the $L^p$metric. If $f$ is measurable, then we define

$$||f||_p = \left(\int_E |f|^p d\mu\right)^{\frac{1}{p}}$$

which can be infinity in the case that $f$ is not $p$-th integrable over $E$. Thus, a measurable function $f \in L^p(E)$ if and only if $||f||_p < \infty$. $||f||_p$ is called the $L^p$-norm of $f$.

Define the $L^p$distance between $f$ and $g$ by

$$d_p(f,g) = ||f-g||_p$$

where $f, g \in L^p(E)$. It can be proved that $d_p$ is a metric on $L^p(E)$ only if $p \geq 1$. Therefore, from now on, we assume that $p \geq 1$, unless stated otherwise.

To see that $L^p(E)$ is a metric space under $d_p$ for $p \geq 1$, we need to show that $d_p$ satisfies the triangle inequality. To this end, we need some elementary facts.

I) *Convex functions.* We say a function $\varphi : (a, b) \to \mathbb{R}$ is convex, if

$$\varphi(\lambda s + (1 - \lambda)t) \leq \lambda\varphi(s) + (1 - \lambda)\varphi(t) \tag{8.1}$$

for all $s, t \in (a, b)$ and $\lambda \in [0, 1]$. The convex property is equivalent to the following: for any $a < s < u < t < b$ we have

$$\frac{\varphi(u) - \varphi(s)}{u - s} \leq \frac{\varphi(t) - \varphi(u)}{t - u} \tag{8.2}$$

by setting $u = \lambda s + (1 - \lambda)t$. It follows that, if $\varphi$ is convex, then the left and right derivatives $\varphi'(t-)$ and $\varphi'(t+)$ of $\varphi$ at any $t \in (a, b)$ must exist, and in fact for every $s \in (a, b)$. The previous inequality implies that

$$\varphi'(s-) = \sup_{a<u<s} \frac{\varphi(u) - \varphi(s)}{u - s}$$

and

$$\varphi'(s+) = \inf_{s<t<b} \frac{\varphi(t) - \varphi(s)}{t - s}.$$

By definition, $\varphi$ is convex in $(a, b)$ if and only if for $a < s < t < u < v < b$ we have

$$\frac{\varphi(t) - \varphi(s)}{t - s} \leq \frac{\varphi(v) - \varphi(u)}{v - u}. \tag{8.3}$$

Thus, both derivatives $t \to \varphi'(t-)$ and $t \to \varphi'(t+)$ are increasing functions on $(a, b)$, and $\varphi'(t-) \leq \varphi'(t+)$ for every $t \in (a, b)$. In particular, if $\varphi$ is convex and differentiable on $(a, b)$, then $\varphi'$ is increasing on $(a, b)$. Of course a convex function is not necessary to be differentiable. We however show that a convex function on $(a, b)$ must be continuous on $(a, b)$.

In fact, using (8.3) again we have

$$\frac{\varphi(t) - \varphi(s)}{t - s} \leq \inf_{b>u>v>t} \frac{\varphi(u) - \varphi(v)}{u - v} \leq \varphi'(u+)$$

and (applying (8.3) with the order $(s, t)$ and $(v, u)$ exchanged)

$$\frac{\varphi(t) - \varphi(s)}{t - s} \geq \sup_{a<v<u<s} \frac{\varphi(u) - \varphi(v)}{u - v} \geq \varphi'(v-)$$

for every $a < v < s < t < u < b$, which follows that

$$|\varphi(t) - \varphi(s)| \leq \max\{|\varphi'(u+)|, |\varphi'(v-)|\}|t - s|$$

for any $a < v < s < t < u < b$. Therefore $\varphi$ is locally Lipschitz continuous, so $\varphi$ must be continuous on $(a, b)$.

On the other hand, if the second derivative $\varphi''$ exists on $(a, b)$ and $\varphi''(t) \geq 0$ for every $t \in (a, b)$ then $\varphi'$ is increasing. By applying MVT to $\varphi'$ we may conclude that 8.2 holds, thus $\varphi$ is convex on $(a, b)$.

A simple example of convex functions is the power function $\varphi(t) = t^p$ for $t > 0$, where $p \geq 1$, then $\varphi''(t) = p(p-1)t^{p-2}$ is non-negative on $(0, \infty)$ so that $t^p$ is convex on $(0, \infty)$. Therefore we have

$$(\lambda s + (1-\lambda)t)^p \leq \lambda s^p + (1-\lambda)t^p \tag{8.4}$$

for all $s, t \geq 0$ and $\lambda \in [0, 1]$.

II) *Hardy's inequality*: If $p > 1$ and $q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$ (such a pair $(p, q)$ of numbers is called a conjugate pair), then

$$st \leq \frac{s^p}{p} + \frac{t^q}{q} \tag{8.5}$$

for all $s, t \geq 0$.

If $t = 0$ then there is nothing to prove. Suppose $t > 0$. The inequality above is equivalent to that

$$\frac{s}{t^{q-1}} \leq \frac{1}{p}\frac{s^p}{t^q} + \frac{1}{q} = \frac{1}{p}\left(\frac{s}{t^{q/p}}\right)^p + \frac{1}{q}.$$

Since $\frac{q}{p} = q - 1$, (8.5) is equivalent to [with $x = \frac{s}{t^{q-1}}$]

$$x \leq \frac{1}{p}x^p + \frac{1}{q} \qquad \text{for all } x > 0. \tag{8.6}$$

Let us prove (8.6). Consider $h(x) = x - \frac{1}{p}x^p - \frac{1}{q}$. Then $h(0) = -\frac{1}{q} < 0$, $h(x) \to -\infty$, and $h'(x) = 1 - x^{p-1}$, thus $x = 1$ is the global maximum, and $h(1) = 0$. Therefore $h(x) \leq 0$ for all $x > 0$, which proves (8.6) and therefore (8.5).

Let us now return to our discussion about $L^p$-space.

**Proposition 8.1** *If $p \geq 1$, and $f, g \in L^p(E)$ then*

$$||f + g||_p \leq ||f||_p + ||g||_p. \tag{8.7}$$

**Proof.** If $||f||_p$ or $||g||_p$ vanishes, then $||f+g||_p = ||f||_p + ||g||_p$, so (8.7) is an equality in this case. Suppose that $||f||_p \neq 0$ and $||g||_p \neq 0$. Apply (8.4) with $s = \frac{|f|}{||f||_p}$ and $t = \frac{|g|}{||g||_p}$ to obtain

$$\left(\lambda\frac{|f|}{||f||_p} + (1-\lambda)\frac{|g|}{||g||_p}\right)^p \leq \lambda\frac{|f|^p}{||f||_p^p} + (1-\lambda)\frac{|g|^p}{||g||_p^p}$$

for every $\lambda \in [0, 1]$. Integrating both sides of the inequality over $E$ we obtain

$$\int_E \left(\frac{\lambda}{||f||_p}|f| + \frac{1-\lambda}{||g||_p}|g|\right)^p d\mu \leq \lambda \int_E \frac{|f|^p}{||f||_p^p} d\mu + (1-\lambda)\int_E \frac{|g|^p}{||g||_p^p} d\mu = 1. \tag{8.8}$$

Choose $\lambda$ such that

$$\frac{\lambda}{||f||_p} = \frac{1-\lambda}{||g||_p}$$

i.e.

$$\lambda = \frac{||f||_p}{||g||_p + ||f||_p}$$

which belongs to $[0,1]$, and plug the value $\lambda$ into (8.8) to deduce that

$$\int_E (|f|+|g|)^p\, d\mu \le (||g||_p + ||f||_p)^p \,.$$

Therefore

$$\int_E |f+g|^p d\mu \le \int_E (|f|+|g|)^p\, d\mu \le (||g||_p + ||f||_p)^p$$

which yields (8.7). ■

**Proposition 8.2** *Let $p \ge 1$. Then $||\cdot||_p$ is a* norm *on $L^p(E)$. That is, the mapping $f \to ||f||_p$ possesses the following three properties:*
*1) $||f||_p = 0$ if and only if $f = 0$ almost everywhere on $E$,*
*2) $||\alpha f||_p = |\alpha|||f||_p$ if $f \in L^p(E)$ and $\alpha$ is a constant,*
*3) Triangle inequality holds:*

$$||f+g||_p \le ||f||_p + ||g||_p.$$

Let us derive another important inequality related to $L^p$-norms.

**Proposition 8.3** *(Hölder's inequality) Let $1 < p, q < \infty$ be a pair of conjugate numbers in the sense that*

$$\frac{1}{p} + \frac{1}{q} = 1$$

*(for example, if $p = 2$, then $q = 2$, and if $p = 3$ then $q = \frac{3}{2}$, and etc.), and let $f \in L^p(E)$ and $g \in L^q(E)$. Then $fg \in L^1(E)$, and*

$$\int_E |fg| d\mu \le ||f||_p ||g||_q \;.$$

**Proof.** (*The proof is not examinable*). Let us now prove Hölder's inequality. If $||f||_p = 0$ or $||g||_q = 0$, then there is nothing to prove, so assume that both $||f||_p \ne 0$ and $||g||_q \ne 0$. Apply (8.5) with

$$s = \frac{|f|}{||f||_p} \quad \text{and } t = \frac{|g|}{||g||_q}$$

and integrate over $E$ to obtain

$$\int_E \frac{|fg|}{||f||_p ||g||_p} \le \frac{1}{p}\int_E \frac{|f|^p}{||f||_p^p} + \frac{1}{q}\int_E \frac{|g|^q}{||g||_q^q} = 1$$

which yields Hölder's inequality. ■

*Example 1.* Let $E$ be measurable with $m(E) < \infty$. Suppose $1 < p_1 < p_2$, and suppose $f$ is measurable. Let $p, q$ be a conjugate pair: $\frac{1}{p} + \frac{1}{q} = 1$, where $1 < p < \infty$. Apply Hölder's inequality to $|f|^{p_1}$ and $g = 1$, to obtain

$$\begin{aligned} \int_E |f|^{p_1} &= \int_E |f|^{p_1} 1 \le \left(\int_E |f|^{p_1 p}\right)^{\frac{1}{p}} \left(\int_E 1^q\right)^{\frac{1}{q}} \\ &= m(E)^{\frac{1}{q}} \left(\int_E |f|^{p_1 p}\right)^{\frac{1}{p}} . \end{aligned} \tag{8.9}$$

Choose $p > 1$ such that $p_1 p = p_2$ which is possible and $p = \frac{p_2}{p_1} > 1$ by assumption that $p_2 > p_1$. Moreover $\frac{1}{p} = \frac{p_1}{p_2}$ and

$$1 = \frac{1}{p} + \frac{1}{q} = \frac{p_1}{p_2} + \frac{1}{q}$$

which yields that

$$\frac{1}{q} = 1 - \frac{p_1}{p_2} = \frac{p_2 - p_1}{p_2} \quad \text{and } q = \frac{p_2}{p_2 - p_1}.$$

Inserting these values into (8.9) we obtain

$$\int_E |f|^{p_1} \leq m(E)^{\frac{p_2-p_1}{p_2}} \left( \int_E |f|^{p_2} \right)^{\frac{p_1}{p_2}}$$

and therefore

$$||f||_{p_1} \leq m(E)^{\frac{p_2-p_1}{p_1 p_2}} ||f||_{p_2}.$$

**Corollary 8.4** *1) If $m(E) < \infty$ and $1 \leq p_1 \leq p_2$, then $L^{p_2}(E) \subset L^{p_1}(E)$.*

*2) If $m(E) = 1$, then $p \to ||f||_p$ is increasing, that is, if $1 \leq p_1 \leq p_2$, then $||f||_{p_1} \leq ||f||_{p_2}$.*

We end this section with some comments about $L^\infty(E)$. A measurable function $f$ on $E$ belongs to $L^\infty(E)$, if there is a null set $A \subset E$ and a constant $M$ such that

$$|f(x)| \leq M \qquad \text{for all } x \in E \setminus A. \tag{8.10}$$

Such a measurable function $f$ is also called *essentially bounded* on $E$. The infinimum of $M$ such that (8.10) holds for some null subset $A$ is called the $L^\infty$-norm of $f$, denoted by $||f||_\infty$. That is

$$||f||_\infty = \inf \left\{ \sup_{x \in E \setminus A} |f(x)| : A \subset E \text{ such that } \mu(A) = 0 \right\}.$$

$||f||_\infty$ is also called the *essential supremum* of $|f|$ over $E$.

Once again, $L^\infty(E)$ is a vector space, and $f \to ||f||_\infty$ is a norm. That is 1) $||f||_\infty \geq 0$, and $||f||_\infty = 0$ if and only if $f = 0$ almost everywhere on $E$; 2) $||\alpha f|| = |\alpha| ||f||_\infty$ for any scalar $\alpha$ and $f \in L^\infty(E)$; and 3) the triangle inequality holds

$$||f + g||_\infty \leq ||f||_\infty + ||g||_\infty.$$

Therefore $L^\infty(E)$ is a metric space with the metric $(f, g) \to ||f - g||_\infty$.

The Hölder inequality applies to $p \in [1, \infty]$:

$$||fg||_1 \leq ||f||_p ||g||_q$$

as long as $\frac{1}{p} + \frac{1}{q} = 1$ (with the convention that $\frac{1}{\infty} = 0$).

The notation $L^\infty(E)$ is justified as we have the following

**Proposition 8.5** *Suppose that $E$ has finite measure (i.e. $\mu(E) < \infty$), and suppose that $f \in L^\infty(E)$, then*

$$||f||_\infty = \lim_{p \to \infty} ||f||_p \ .$$

**Proof.** If $||f||_\infty = 0$, then there is nothing to prove as $||f||_p = 0$ for all $p$. Assume that $||f||_\infty > 0$. Then $\mu(E) > 0$. For every $\varepsilon > 0$ there is a null set $A \subset E$ such that $|f(x)| \leq ||f||_\infty + \varepsilon$ for all $x \in E \setminus A$, hence

$$||f||_p \leq (||f||_\infty + \varepsilon)\,(\mu(E))^{\frac{1}{p}}$$

since $(\mu(E))^{\frac{1}{p}} \to 1$ as $p \to \infty$,

$$\limsup ||f||_p \leq ||f||_\infty + \varepsilon,$$

and since $\varepsilon > 0$ is arbitrary, we must have $\limsup ||f||_p \leq ||f||_\infty$. On the other hand, for every $\varepsilon \in (0, ||f||_\infty)$, there is a subset $B \subset E$ with $\mu(B) > 0$ and $|f(x)| \geq ||f||_\infty - \varepsilon$ for all $x \in B$. Therefore

$$||f||_p \geq \left(\int_B |f|^p\right)^{\frac{1}{p}} \geq (||f||_\infty - \varepsilon)\,\mu(B)^{\frac{1}{p}} \to ||f||_\infty - \varepsilon$$

as $p \to \infty$, which yields that $\liminf ||f||_p \geq ||f||_\infty - \varepsilon$. Letting $\varepsilon \downarrow 0$ to obtain $\liminf ||f||_p \geq ||f||_\infty$. This completes the proof. ■

## 8.2 Convergence in measure

Recall that we are working with a measure space $(E, \mathcal{F}, \mu)$ which is complete in the sense that if $\mu^*(A) = 0$ where $A \subseteq E$, then $A \in \mathcal{F}$. Our basic example is that $E$ is a Lebesgue measurable subset of $\mathbb{R}^n$, $\mathcal{F} = \mathcal{M}_{\text{Leb}}(E)$, and $\mu = m$ is the Lebesgue measure.

**Definition 8.6** *$f_n \to f$ almost everywhere on $E$ as $n \to \infty$, if there is a null subset $A \subset E$, such that $f_n(x) \to f(x)$ for every $x \in E \setminus A$.*

Suppose $f_n, f : E \to \mathbb{C}$ are measurable, then, $x \in \{f_n \to f\}$, if and only if every $\varepsilon > 0$, there is $N$ such that $x \in \{|f_n - f| \leq \varepsilon\}$ for all $n \geq N$, that is, if and only if for every $\varepsilon > 0$

$$x \in \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \{|f_n - f| \leq \varepsilon\}.$$

Thus, $f_n \to f$ almost everywhere on $E$ if and only if for every $\varepsilon > 0$,

$$\mu\left(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} \{|f_n - f| > \varepsilon\}\right) = 0.$$

In particular, if $\mu(E) < \infty$, $f_n, f : E \to \mathbb{R}$ are measurable, then $f_n \to f$ almost everywhere on $E$ if and only if for every $\varepsilon > 0$

$$\lim_{N\to\infty} \mu\left(\bigcup_{n=N}^{\infty} \{|f_j - f| > \varepsilon\}\right) = 0.$$

Therefore, if $\mu(E) < \infty$ and $f_n \to f$ almost everywhere on $E$, then

$$\mu\left[|f_n - f| > \varepsilon\right] \to 0 \quad \text{as } n \to \infty.$$

**Definition 8.7** *Let $E \subset \mathbb{R}$ be a measurable subset, and $f_n, f : E \to \mathbb{R}$ be measurable. Then $f_n \to f$* in measure *if for every $\varepsilon > 0$*

$$\mu\left(\{|f_n - f| > \varepsilon\}\right) \to 0 \tag{8.11}$$

*as $n \to \infty$, where*

$$\{|f_n - f| > \varepsilon\} = \{x \in E : |f_n(x) - f(x)| > \varepsilon\}.$$

If $\mu(E) < \infty$, $f_n \to f$ almost everywhere on $E$ implies that $f_n \to f$ in measure. Conversely, we have the following

**Proposition 8.8** *(F. Riesz) If $f_n \to f$ in measure, then there is a sub-sequence $(f_{n_k})$ such that $f_{n_k} \to f$ almost surely.*

**Proof.** (*The proof is not examinable*) Since $f_n \to f$ in measure, by definition, for every $k = 1, 2, \cdots$, there is $n_k$ such that

$$\mu\left(\left\{|f_n - f| > \frac{1}{2^k}\right\}\right) < \frac{1}{2^k} \qquad \forall n \geq n_k.$$

We may choose $n_k$ so that $n_1 < n_2 < \cdots$.

We next show that $f_{n_k} \to f$ almost everywhere on $E$. To this end, consider

$$E_k = \bigcap_{j=k}^{\infty} \left\{|f_{n_j} - f| \leq \frac{1}{2^j}\right\}$$

where $k = 1, 2, \cdots$, and set $G = \bigcup_{k=1}^{\infty} E_k$. Clearly, if $x \in G$ then

$$|f_{n_j}(x) - f(x)| \leq \frac{1}{2^j} \quad \text{for all } j \geq k$$

for some $k$. Therefore $G \subset \{f_{n_k} \to f\}$. On the other hand

$$E \setminus G = \bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} \left\{|f_{n_j} - f| > \frac{1}{2^j}\right\}$$

so that

$$\begin{aligned} \mu(E \setminus G) &\leq \mu\left(\bigcup_{j=k}^{\infty} \left\{|f_{n_j} - f| > \frac{1}{2^j}\right\}\right) \\ &\leq \sum_{j=k}^{\infty} \mu\left(\left\{|f_{n_j} - f| > \frac{1}{2^j}\right\}\right) \leq \sum_{j=k}^{\infty} \frac{1}{2^j} \end{aligned}$$

for any $k$. Since $\sum_{j=k}^{\infty} \frac{1}{2^j} \to 0$ as $k \to \infty$, therefore $\mu(E \setminus G) = 0$. The proof is complete. ■

## 8.3 Convergence in $L^p$-space

Recall that for $p \geq 1$, $L^p(E)$ is a vector space, and $d_p(f,g) = ||f-g||_p$ is a metric on $L^p(E)$.

**Definition 8.9** *Let $p \geq 1$. We say $f_n \to f$ in $L^p$-norm if $||f_n - f||_p \to 0$ as $n \to \infty$.*

**Proposition 8.10** *If $f_n \to f$ in $L^p$-norm, then $f_n \to f$ in measure, and therefore there is a sub-sequence $(f_{n_k})$ such that $f_{n_k} \to f$ almost everywhere.*

**Proof.** For any $\varepsilon > 0$ we have

$$\begin{aligned}\mu(\{|f_n - f| &> \varepsilon\}) = \int_E 1_{\{|f_n - f|>\varepsilon\}} d\mu \leq \int_E \frac{|f_n - f|^p}{\varepsilon^p} 1_{\{|f_n-f|>\varepsilon\}} d\mu \\ &\leq \frac{1}{\varepsilon^p}\int_E |f_n - f|^p d\mu = \frac{1}{\varepsilon^p}||f_n - f||_p^p\end{aligned}$$

which goes to zero as $n \to \infty$. The second claim then follows from Proposition 8.8. ■

**Theorem 8.11** *Let $p \geq 1$. $L^p(E)$ with the metric $d_p(f,g) = ||f-g||_p$ is a complete metric space. That is, if $(f_n)$ is a Cauchy sequence in $L^p(E)$: $f_n \in L^p(E)$ and*

$$||f_n - f_m||_p \to 0 \quad \text{as } n, m \to \infty$$

*then*

*1) there is a sub-sequence $(f_{n_k})$ and $f \in L^p(E)$ such that $f_{n_k} \to f$ almost everywhere on $E$, and*

*2) $||f_n - f||_p \to 0$ as $n \to \infty$.*

Therefore $L^p(E)$ equipped with the norm $||\cdot||_p$ is a *Banach space.*

**Proof.** (*The proof is not examinable*) 1) The proof is the modification of that of Proposition 8.8. Since $||f_n - f_m||_p \to 0$ so for every $k = 1, 2, \cdots$, there is $n_k$ such that

$$||f_n - f_m||_p^p \leq \frac{1}{2^k}\left(\frac{1}{2^k}\right)^p \quad \text{for any } n, m \geq n_k.$$

We can choose $n_k$ such that $n_k$ is strictly increasing. Then

$$\begin{aligned}\mu\left(\left\{|f_n - f_m| \geq \frac{1}{2^k}\right\}\right) &= \int_E 1_{\{|f_n - f_m| \geq \frac{1}{2^k}\}} d\mu \leq \int_E \frac{|f_n - f_m|^p}{\left(\frac{1}{2^k}\right)^p} d\mu \\ &= \frac{||f_n - f_m||_p^p}{\left(\frac{1}{2^k}\right)^p} \leq \frac{1}{2^k} \qquad \forall n, m \geq n_k. \end{aligned} \tag{8.12}$$

Let

$$E_k = \bigcap_{j=k}^{\infty}\left\{|f_{n_{j+1}} - f_{n_j}| \leq \frac{1}{2^j}\right\}, \quad G = \bigcup_{k=1}^{\infty} E_k$$

Then if $x \in G$, $|f_{n_{i+1}}(x) - f_{n_i}(x)| \leq \frac{1}{2^i}$ for all $i \geq k$ for some $k$, so that

$$|f_{n_i}(x) - f_{n_j}(x)| \leq \sum_{i=j}^{\infty}|f_{n_{i+1}}(x) - f_{n_i}(x)| \leq \sum_{i=j}^{\infty}\frac{1}{2^i} = \frac{1}{2^{j-1}} \quad \forall i \geq j \geq k$$

and therefore $(f_{n_k}(x))$ converges to some $f(x)$ for every $x \in G$. On the other hand

$$E \setminus G = \bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} \left\{ |f_{n_{j+1}} - f_{n_j}| \geq \frac{1}{2^j} \right\}$$

so that, by (8.12)

$$\mu(E \setminus G) \leq \sum_{j=k}^{\infty} \mu \left( \left\{ |f_{n_{j+1}} - f_{n_j}| \geq \frac{1}{2^j} \right\} \right) \leq \sum_{j=k}^{\infty} \frac{1}{2^j} = \frac{1}{2^{k-1}}$$

for any $k = 1, 2, \cdots$, hence $\mu(E \setminus G) = 0$. Therefore $(f_{n_k})$ converges to $f$ almost everywhere.

We next show that $f \in L^p(E)$ and $||f_n - f||_p \to 0$ as $n \to \infty$. For every $\varepsilon > 0$ there is $N$ such that $||f_n - f_m||_p < \varepsilon$ for all $n, m \geq N$. Therefore

$$\int_E |f_n - f_{n_k}|^p d\mu < \varepsilon^p \text{ for any } k, n \geq N.$$

Since $|f_n - f_{n_k}| \to |f_n - f|$ almost everywhere on $E$, so by Fatou's Lemma

$$\int_E |f_n - f|^p d\mu \leq \underline{\lim}_{k \to \infty} \int_E |f_n - f_{n_k}|^p d\mu \leq \varepsilon^p \qquad \forall n \geq N,$$

therefore $f_n - f \in L^p(E)$ and $||f_n - f||_p \leq \varepsilon$. By definition $f_n \to f$ in $L^p(E)$ and $f \in L^p(E)$. ■

# 9 Appendix 1 – From outer measures to measures

*The material below is not on the syllabus of A4 Paper Integration, which covers the essential steps in the construction of Lebesgue measures.*

**Theorem 9.1** *Let $\mathcal{R}$ be a* ring *of some subsets of $\Omega$, that is, $\mathcal{R}$ contains the empty set $\emptyset$ and is closed under $\cup$ and $\cap$. Suppose there is a sequence $E_n$ ($n = 1, 2, \cdots$) such that $\bigcup_{n=1}^{\infty} E_n = \Omega$. Suppose $\mu : \mathcal{R} \to [0, \infty]$ satisfies the following conditions: (1) $\mu(\emptyset) = 0$; (2) $\mu(A) \leq \mu(B)$ for $A, B \in \mathcal{R}$ and $A \subset B$; and (3) $\mu$ is finitely additive: if $A_i$ are disjoint, $A_i \in \mathcal{R}$, where $i = 1, \cdots, n$, then*

$$\mu \left[ \bigcup_{i=1}^{n} A_i \right] = \sum_{i=1}^{n} \mu(A_i). \tag{9.1}$$

*Let $\mu^*$ be the outer measure defined by $\mu$:*

$$\mu^*(E) = \inf \left\{ \sum_{i=1}^{\infty} \mu(E_i) : E_i \in \mathcal{R} \text{ such that } \bigcup_{i=1}^{\infty} E_i \supset E \right\} \tag{9.2}$$

*for $E \subset \Omega$. Let $\mathcal{G}^m$ be the $\sigma$-algebra of $\mu^*$-measurable subsets. Then $\mathcal{R} \subset \mathcal{G}^m$.*

*If in addition we assume that $\mu^*(A) = \mu(A)$ for every $A \in \mathcal{R}$, then $\mu^*$ is an extension of $\mu$ (to a measure on $\mathcal{G}^m$).*

[But, be careful, $\mu^*$ restricted on $\mathcal{G}^m$ is not necessary an extension of $\mu$. In fact, $\mu^*$ restricted on the ring $\mathcal{R}$ does not need to coincide with $\mu$!]

**Proof.** *[The proof is not examinable].* If $E \in \mathcal{R}$, in order to show that $E$ is $\mu^*$-measurable, we show that

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

for every $F \subseteq \mathbb{R}$. It is trivial if $\mu^*(F) = \infty$, so we consider the case that $\mu^*(F) < \infty$. For every $\varepsilon > 0$, there is a sequence $\{E_i : i = 1, 2, \cdots\}$, where $E_i \in \mathcal{R}$, such that $\cup_{i=1}^{\infty} E_i \supseteq F$ and

$$\sum_{i=1}^{\infty} \mu(E_i) \leq \mu^*(F) + \varepsilon.$$

Let $A_i = E_i \cap E$ and $B_i = E_i \cap E^c$. Then $A_i, B_i \in \mathcal{R}$, $A_i \cup B_i = E_i$ and $A_i \cap B_i = \emptyset$, so that, by finite additivity of $\mu$ over the ring $\mathcal{R}$, we have

$$\mu(E_i) = \mu(A_i) + \mu(B_i).$$

Since $\cup_i A_i \supseteq F \cap E$ and $\cup_i B_i \supseteq F \cap E^c$, so that

$$\mu^*(F \cap E) \leq \sum_{i=1}^{\infty} \mu(A_i), \text{ and } \mu^*(F \cap E^c) \leq \sum_{i=1}^{\infty} \mu(B_i).$$

It follows that

$$\mu^*(F \cap E) + \mu^*(F \cap E^c) \leq \sum_{i=1}^{\infty} \mu(A_i) + \sum_{i=1}^{\infty} \mu(B_i) \leq m^*(F) + \varepsilon.$$

Letting $\varepsilon \downarrow 0$ to obtain

$$\mu^*(F \cap E) + \mu^*(F \cap E^c) \leq \mu^*(F)$$

which shows that $E \in \mathcal{R}$ is $\mu^*$-measurable. ∎

This theorem will be used to give a complete description of the $\sigma$-algebra of Lebesgue measurable subsets of $\mathbb{R}$ in the next sub-section. Let us prove the extension theorem of Caratheodory's.

**Theorem 9.2** *Let $\mathcal{R}$ be a* ring *of some subsets of $\Omega$, such that there is a sequence $E_n$ ($n = 1, 2, \cdots$) such that $\bigcup_{n=1}^{\infty} E_n = \Omega$. Let $\mu : \mathcal{R} \to [0, \infty]$ be a measure on the ring $\mathcal{R}$. That is, $\mu(\emptyset) = 0$; $\mu(A) \leq \mu(B)$ for $A, B \in \mathcal{R}$ and $A \subset B$; and $\mu$ is conutably additive: if $A_i$ are disjoint, $A_i \in \mathcal{R}$, where $i = 1, 2, \cdots$, such that $\bigcup_{i=1}^{\infty} A_i \in \mathcal{R}$ then*

$$\mu\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} \mu(A_i). \tag{9.3}$$

*Let $\mu^*$ be the outer measure associated with $\mu$ defined by (9.2), and let $\mathcal{G}^m$ be the $\sigma$-algebra of $\mu^*$-measurable subsets. Then $\mathcal{R} \subset \mathcal{G}^m$ and$\mu^*(A) = \mu(A)$ for every $A \in \mathcal{R}$. Therefore $\mu^*$ is an extension of $\mu$ (to a measure on $\mathcal{G}^m$).*

**Proof.** *[The proof is not examinable].* According to Theorem 9.1, we only need to show that $\mu^*(E) = \mu(E)$ for every $E \in \mathcal{R}$. By definition $\mu^*(E) \le \mu(E)$. Let us prove that $\mu(E) \le \mu^*(E)$. If $\mu^*(E) = \infty$, then there is nothing to prove, so we consider the case that $\mu^*(E) < \infty$. For every $\varepsilon > 0$, there is a countable cover of $E$, $\{E_i\}$, where $E_i \in \mathcal{R}$ for $i = 1, 2, \cdots$, such that

$$\sum_{i=1}^{\infty} \mu(E_i) < \mu^*(E) + \varepsilon.$$

Let $A_1 = E \cap E_1$, $A_i = E \bigcap \left(E_i \setminus \cup_{k=1}^{i-1} E_k\right)$ for $i \ge 2$. Then $A_i \in \mathcal{R}$ are disjoint, and $\bigcup_{i=1}^{\infty} A_i = E$, thus, since $\mu$ is a measure on the ring $\mathcal{R}$, we have

$$\mu(E) = \sum_{i=1}^{\infty} \mu(A_i) \le \sum_{i=1}^{\infty} \mu(E_i) < \mu^*(E) + \varepsilon$$

for every $\varepsilon > 0$. Letting $\varepsilon \downarrow 0$ to obtain $\mu(E) \le \mu^*(E)$. The proof is completed. ■

Carathéodory's extension theorem, more explicitly Theorem 9.1, may be applied to the construction of the Lebesgue measure $m$ on $\mathbb{R}$. Recall that the outer measure $m^*$ is defined on $\mathcal{P}(\mathbb{R})$ of all subsets of $\mathbb{R}$, so the $m^*$-measurable subsets, called Lebesgue measurable in this case, or simply measurable in this course if no confusion may arise, form a $\sigma$-algebra $\mathcal{P}(\mathbb{R})^m$, which will be denoted as $\mathcal{M}_{\text{Leb}}$ or $\mathcal{M}_{\text{Leb}}(\mathbb{R})$. Then the outer measure $m^*$ restricted on $\mathcal{M}_{\text{Leb}}$ is a measure, the triple $(\mathbb{R}, \mathcal{M}_{\text{Leb}}, m^*)$ is called the *Lebesgue measure space.*

Our next task is to describe the $\sigma$-algebra $\mathcal{M}_{\text{Leb}}$ of Lebesgue measurable sets on $\mathbb{R}$.

The following lemma is the *key technical fact* used in the construction of Lebesgue's measure $m$ on $\mathbb{R}$.

**Lemma 9.3** *If $A = \bigcup_{i=1}^{n} J_i$, where $J_i \in \mathcal{C}$ are disjoint, then $m^*(A) = \sum_{i=1}^{n} |J_i|$.*

**Proof.** (*The proof is not examinable.*) Since $\{J_i\}$ is a finite cover of $A$, and $J_i \in \mathcal{C}$, so that $m^*(A) \le \sum_{i=1}^{n} |J_i|$. We first prove that $m^*(J) = m(J) = |J|$ for every $J = (a, b] \in \mathcal{C}$. By definition, $m^*(J) \le |J|$. On the other hand, for every $\varepsilon > 0$, there is a countable cover $\{I_k = (s_k, t_k] : k = 1, 2, \cdots\}$ of $(a, b]$, such that $m^*(J) \ge \sum_{k=1}^{\infty} |I_k| - \varepsilon$. Since $[a, b]$ is closed and bounded, and $\left\{(s_k - \frac{\varepsilon}{2^{k+1}}, t_k + \frac{\varepsilon}{2^{k+1}}) : k = 1, 2, \cdots\right\}$ is an open cover of $[a, b]$, there is a finite sub-cover $\{(s_{k_i} - \frac{\varepsilon}{2^{k_i+1}}, t_{k_i} + \frac{\varepsilon}{2^{k_i+1}}) : i = 1, \cdots, N\}$ for some $N$. [This is a result from Analysis II]. Let us write $a_i = s_{k_i} - \frac{\varepsilon}{2^{k_i+1}}$ and $b_i = t_{k_i} + \frac{\varepsilon}{2^{k_i+1}}$ $(i = 1, \cdots, N)$ for simplicity, and, if necessary, rearrange them in the following order

$$a_1 \le a_2 \le \cdots \le a_N.$$

Then we must have $b_k \ge a_{k+1}$ (for any $k = 1, 2, \cdots, N-1$), $a_1 \le a$ and $b_N \ge b$, so that $\sum_{i=1}^{N}(b_i - a_i) \ge b - a$. It follows that

$$\sum_{k=1}^{\infty} |I_k| \ge \sum_{i=1}^{N} |I_{k_i}| = \sum_{i=1}^{N}(b_i - a_i) - 2\sum_{k=1}^{\infty} \frac{\varepsilon}{2^{k+1}} \ge b - a - \varepsilon.$$

Therefore $m^*(J) \ge (b - a) - \varepsilon$. Since $\varepsilon$ is arbitrary, $m^*(J) \ge b - a$.

Let us now consider the general case. For every $\varepsilon > 0$ there is a sequence $I_k \in \mathcal{C}$ of intervals such that $\cup_{k=1}^{\infty} I_k \supseteq A$ and

$$m^*(A) \leq \sum_{k=1}^{\infty} |I_k| \leq m^*(A) + \varepsilon.$$

Note that $I_k \cap J_i \subset I_k$ and $I_k \cap J_i \in \mathcal{C}$, so that $\{I_k \cap J_i : i = 1, \cdots, n\}$ are disjoint sub-intervals of $I_k$ for fixed $k$, hence $\sum_{i=1}^{n} |I_k \cap J_i| \leq |I_k|$. On the other hand

$$\begin{aligned} |J_i| &= m^*(J_i) = m^*\left(\cup_{k=1}^{\infty} I_k \cap J_i\right) \\ &\leq \sum_{k=1}^{\infty} m^*(I_k \cap J_i) = \sum_{k=1}^{\infty} |I_k \cap J_i| \end{aligned}$$

and therefore

$$\begin{aligned} \sum_{i=1}^{n} |J_i| &\leq \sum_{i=1}^{n} \sum_{k=1}^{\infty} |I_k \cap J_i| = \sum_{k=1}^{\infty} \sum_{i=1}^{n} |I_k \cap J_i| \\ &\leq \sum_{k=1}^{\infty} |I_k| \leq m^*(A) + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we may conclude that $\sum_{i=1}^{n} |J_i| \leq m^*(A)$. Hence $m^*(A) = \sum_{i=1}^{n} |J_i|$. ■

Let $\mathcal{R}$ be the collection of all subsets of $\mathbb{R}$ which can be written as a disjoint union of finite many elements in the $\pi$-system $\mathcal{C}$, i.e. $E \in \mathcal{R}$ if $E = \cup_{i=1}^{n} J_i$ where $J_i \in \mathcal{C}$ are disjoint for some $m$. Then $\mathcal{R}$ is a *ring* of subsets of $\mathbb{R}$, that is, $\mathcal{R}$ contains the empty $\varnothing$ and is closed under $\cup$ and $\cap$.

If $E = \cup_{i=1}^{n} J_i \in \mathcal{R}$, where $J_i \in \mathcal{C}$ are disjoint, so that $m^*(E) = \sum_{i=1}^{n} |J_i|$ does not depend on the representation of $E$ as finite disjoint union of elements in $\mathcal{C}$. $m^* : \mathcal{R} \to [0, \infty]$ possesses the following properties:

1) $m^*(\varnothing) = 0$,

2) $m^*$ is *finitely additive on the ring* $\mathcal{R}$: if $E_1, \cdots, E_n \in \mathcal{R}$ are disjoint, then $m^*(\cup_{j=1}^{n} E_j) = \sum_{j=1}^{n} m^*(E_j)$,

3) $m^*$ is countably sub-additive.

In fact 1) and 3) hold as $m^*$ is an outer measure on $\mathcal{P}(\mathbb{R})$. 2) follows from Lemma 9.3 immediately. In fact, suppose $E_j = \cup_{k=1}^{n_j} I_k^{(j)}$ where $I_k^{(j)} \in \mathcal{C}$ are disjoint, so that

$$\bigcup_{j=1}^{n} E_j = \bigcup_{j=1}^{n} \bigcup_{k=1}^{n_j} I_k^{(j)}$$

is a disjoint decomposition of $\cup_{j=1}^{n} E_j$ and therefore, by Lemma 9.3,

$$m^*\left(\bigcup_{j=1}^{n} E_j\right) = \sum_{j=1}^{n} \sum_{k=1}^{n_j} |I_k^{(j)}| = \sum_{j=1}^{n} m^*(E_j)$$

which proves 2).

With the help of these facts about the outer measure $m^*$, we have the following

**Lemma 9.4** *The Lebesgue outer measure*

$$m^*(E) = \inf \left\{ \sum_{i=1}^{\infty} m^*(E_i) : \text{ where } E_i \in \mathcal{R} \text{ such that } \bigcup_{i=1}^{\infty} E_i \supset E \right\} \tag{9.4}$$

*for any subset $E \subseteq \mathbb{R}$, where the inf runs over all countable covers $\{E_i : i = 1, 2, \cdots\} \subset \mathcal{R}$ of $E$.*

**Proof.** For $E \subseteq \mathbb{R}$, let us use $\varphi(E)$ to denote the right-hand side of (9.4). We want to show that $\varphi(E) = m^*(E)$. Since the $\pi$-system $\mathcal{C} \subset \mathcal{R}$, so that, by definition of the outer measure $m^*(E)$, we have $\varphi(E) \leq m^*(E)$. Let us prove that $m^*(E) \leq \varphi(E)$. If $\varphi(E) = \infty$, then there is nothing to prove, so we consider the case that $\varphi(E) < \infty$. For every $\varepsilon > 0$, there is a countable cover of $E$, $\{E_i\}$, where $E_i \in \mathcal{R}$ for $i = 1, 2, \cdots$, such that

$$\sum_{i=1}^{\infty} m^*(E_i) - \varepsilon < \varphi(E) \leq \sum_{i=1}^{\infty} m^*(E_i).$$

Since $E_i \in \mathcal{R}$, $E_i = \sum_{j=1}^{n_i} I_j^{(i)}$ where $I_j^{(i)} \in \mathcal{C}$ and are disjoint for each $i$. Putting all these together $\{I_j^{(i)}\}$ with two indices $(i, j)$, which is again countable, and thus forms a countable cover of $E$. By Lemma 9.3 $m^*(E_i) = \sum_{j=1}^{n_i} |I_j^{(i)}|$ for each $i$, so that

$$\sum_{i=1}^{\infty} m^*(E_i) = \sum_{i,j} |I_j^{(i)}|.$$

According to definition of $m^*(E)$, we have

$$m^*(E) \leq \sum_{i,j} |I_j^{(i)}| = \sum_{i=1}^{\infty} m^*(E_i) < \varphi(E) + \varepsilon$$

for every $\varepsilon > 0$. Letting $\varepsilon \downarrow 0$ to obtain $m^*(E) = \varphi(E)$, so the proof is completed. ■

We are now in a position to show the following main structure theorem for the Lebesgue measure space.

**Theorem 9.5** *$\mathcal{R} \subseteq \mathcal{M}_{Leb}$, and therefore $\sigma\{\mathcal{R}\} \subseteq \mathcal{M}_{Leb}$ .*

This follows directly from Theorem 9.1.

We are now in a position to describe the structure of the $\sigma$-algebra of $\mathcal{M}_{\text{Leb}}$.

The $\sigma$-algebra $\sigma\{\mathcal{R}\}$ over $\mathbb{R}$ generated by the ring $\mathcal{R}$ (or equivalently by the $\pi$-system $\mathcal{C}$, $\sigma\{\mathcal{C}\}$) is the Borel $\sigma$-algebra $\mathcal{B}(\mathbb{R})$ (also denoted by $\mathcal{M}_{\text{Bor}}$ for good reason). Any interval $J \in \mathcal{B}(\mathbb{R})$, so are the open and closed subsets of $\mathbb{R}$. Any subset in $\mathcal{B}(\mathbb{R})$ is called Borel measurable. Therefore Borel measurable sets are Lebesgue measurable, open subsets and closed subsets are Borel measurable, so are Lebesgue measurable.

If $J$ is an interval then $J$ is Borel measurable (so must be Lebesgue measurable), and $m(J) = |J|$ the length of $J$. In fact, if $J$ is infinite interval, then there is a number $a$ such that $(a, \infty)$ or $(-\infty, a] \subseteq J$, so that $m(J) \geq m((a, a + n])$ (or $m((a - n, a]))$, thus $m(J) \geq n$ for every

$n \in \mathbb{N}$. Therefore $m(J) = \infty$. Suppose now $J$ is an interval with two ends $a \leq b$. Then $(a+\varepsilon, b-\varepsilon] \subseteq J \subseteq (a-\varepsilon, b+\varepsilon]$, so that

$$b - a - 2\varepsilon = m((a+\varepsilon, b-\varepsilon]) \leq m(J) \leq m((a-\varepsilon, b+\varepsilon]) = b - a + 2\varepsilon$$

for every $\varepsilon > 0$. Letting $\varepsilon \downarrow 0$ to obtain $m(J) = b - a$.

Therefore

$$m^*(E) = \inf \left\{ \sum_{k=1}^{\infty} |J_k| : \text{ where } J_k \text{ are intervals such that } \cup_{k=1}^{\infty} J_k \supseteq E \right\}$$

for every subset $E \subseteq \mathbb{R}$.

We can show that any open set $U$ of $\mathbb{R}$ can be written as a union of at most countable many *disjoint* open intervals

$$U = \cup_{j=1}^{\infty} (a_j, b_j).$$

By the countable additivity we have

$$m(U) = \sum_{j=1}^{\infty} (b_j - a_j) \ .$$

Therefore, Lebesgue measure $m$ is really an extension of the notion of length to measurable sets.

As another application of Theorem 4.3 and Theorem 9.2, we may build a *complete* measure space from a given measure space.

**Theorem 9.6** *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, and let $\mu^*$ be the outer measure associated with $\mu$ defined by*

$$\mu^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu(E_i) : E_i \in \mathcal{F} \ s.t. \ \cup_{i=1}^{\infty} E_i \supset E \right\}$$

*for every $E \subset \Omega$. Let $\mathcal{F}^m$ be the $\sigma$-algebra on $\Omega$ consisting of all $\mu^*$-measurable subsets. Then, according to Theorem 4.3, $(\Omega, \mathcal{F}^m, \mu^*)$ is a measure space, and according to Theorem 9.2, $\mathcal{F} \subseteq \mathcal{F}^m$ and $\mu = \mu^*$ on $\mathcal{F}$. Moreover,$\mathcal{F}^m$ contains all $\mu^*$-null sets.*

The measure space $(\Omega, \mathcal{F}^m, \mu^*)$ is called the Carathéodory extension of $(\Omega, \mathcal{F}, \mu)$.

Thanks to Theorem 9.6, given a measure space $(\Omega, \mathcal{F}, \mu)$, we naturally extend it (in many applications, without further comments) to *its completion* $(\Omega, \mathcal{F}^\mu, \mu^*)$, where $\mathcal{F}^\mu = \sigma\{\mathcal{F}, \mathcal{N}\}$ where $\mathcal{N} \subset \mathcal{F}^*$ the collection of all $\mu^*$-null sets. Then $\mathcal{F}^\mu \subseteq \mathcal{F}^m$ which contains all $\mu^*$-null subsets.

**Definition 9.7** *A measure space $(\Omega, \mathcal{F}, \mu)$ is called* complete*, if for every $A \in \mathcal{F}$ with $\mu(A) = 0$, then any subset of $A$ also belongs to $\mathcal{F}$.*

According to definition, if $(\Omega, \mathcal{F}, \mu)$ is a measure space, then its completion $(\Omega, \mathcal{F}^\mu, \mu^*)$ is a complete measure space. If $E$ is Lebesgue measurable, then the Lebesgue space $(E, \mathcal{M}_{\text{Leb}}, m)$ is complete.

# 10 Appendix 2 – Further results on Riemann integrals

[*This part is not examinable in paper A4 Integration.*] The section is devoted to the proof of an interesting result about Riemann integrals, by means of Lebesgue's measure.

The Riemann integral of a function $f : [a, b] \to \mathbb{R}$, in most text books, is defined to be the limit of Riemann sums

$$\lim_{|D|\to 0} \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1})$$

where the limit runs over all possible finite partitions $D : a = x_0 < x_1 < \cdots < x_n = b$, and $\xi_j \in [x_{j-1}, x_j]$, where $|D| = \max_{j=1,\cdots,n}(x_j - x_{j-1})$. $f$ is *Riemann integrable* if the above limit exists, and its limit is denoted by $\int_a^b f(x)dx$. While the limiting procedure is of course not in the ordinary sense as for sequences, but the idea of limits for functions and sequences can be adopted to Riemann sums directly. More precisely, $f$ is Riemann integrable on $[a, b]$ if there is a number $I$, for every $\varepsilon > 0$, there is $\delta > 0$, such that for whatever finite partition $D$ of $[a, b]$ and for whatever choices of $\xi_j \in [x_{j-1}, x_j]$, as long as $|D| < \delta$, we have

$$\left| \sum_j f(\xi_j)(x_j - x_{j-1}) - I \right| < \varepsilon.$$

The limit $I$ is called the Riemann integral of $f$ on $[a, b]$, denoted as $\int_a^b f(x)dx$.

**Lemma 10.1** *If $f : [a, b] \to \mathbb{R}$ is Riemann integrable on $[a, b]$, then $f$ is bounded on $[a, b]$.*

**Proof.** We can prove this by contradiction. Suppose otherwise, that is, $f$ were unbounded on $[a, b]$. Then there is a sequence $\{p_n\}$ in $[a, b]$ such that $f(p_n) \to \infty$. Since $[a, b]$ is a closed interval, without losing generality, we may assume that $\{p_n\}$ is convergent, and $p_n \to p$. Of course $p \in [a, b]$. Since $f$ is Riemann integrable on $[a, b]$, there is $\delta > 0$ such that

$$\left| \sum_j f(\xi_j)(x_j - x_{j-1}) - I \right| \leq 1$$

for any finite partition $D = \{a = x_0 < \cdots < x_n = b\}$ such that $|D| < \delta$ and for any choices of $\xi_j \in [x_{j-1}, x_j]$. Apply this to the equal partition $x_{N,j} = a + \frac{j}{N}(b - a)$ for every natural number $N$ such that $\frac{1}{N}(b - a) < \delta$, where $j = 0,\cdots, N$. Then

$$\left| \sum_{j=1}^{N} f(\xi_j)\frac{b-a}{N} - I \right| \leq 1 \tag{10.1}$$

for any $\xi_j \in [x_{N,j-1}, x_{N,j}]$. Let us fix such $N$ such that $\frac{1}{N}(b - a) < \delta$. Certainly $p$ belongs to (at most two) some interval $[x_{N,j_0-1}, x_{N,j_0}]$. Since $p_n \to p$, so there are $N_1$ and $j_0$ (at least one, at most two) such that there is a sub-sequence $p_{n_l} \in [x_{N,j_0-1}, x_{N,j_0}]$ for all $l$. Let us choose $\xi_j \in [x_{N,j-1}, x_{N,j}]$ for $j \neq j_0$ and fix them, and choose $\xi_{j_0} = p_l$ in Inequality (10.1) to obtain

$$\left| \sum_{j\neq j_0}^{N} f(\xi_j)\frac{b-a}{N} + \frac{b-a}{N} f(p_{n_l}) - I \right| \leq 1.$$

Letting $l \uparrow \infty$ to obtain $\infty \leq 1$ which is a contradiction. ■

Suppose $f : [a, b] \to \mathbb{R}$ is *bounded*, then the following limit exists:

$$S = \lim_{|D| \to 0} \sum_{j=1}^{n} M_j (x_j - x_{j-1})$$

where again $D$ runs through all possible finite partition $a = t_0 < t_1 < \cdots < t_n = b$, and $M_j = \sup_{[x_{j-1}, x_j]} f$ the supremum of $f$ over $[x_{j-1}, x_j]$. That is, there is a number $S$ such that for every $\varepsilon > 0$, there is $\delta > 0$ such that

$$\left| \sum_j M_j (x_j - x_{j-1}) - S \right| < \varepsilon$$

as long as the finite partition $D = \{a = x_0 < \cdots < x_n = b\}$ satisfying that $|D| < \delta$. This limit $S$ is called the upper integral of $f$ over $[a, b]$, denoted by $\overline{\int}_a^b f(x)dx$. Similarly, by replacing the supremum $M_j$ by $m_j = \inf_{[x_{j-1}, x_j]} f$ we may define the lower integral $\underline{\int}_a^b f(x)dx$. Then, $f$ is Riemann integrable on $[a, b]$, if and only if its upper integral $\overline{\int}_a^b f(x)dx$ coincides with its lower integral $\underline{\int}_a^b f(x)dx$, and the common value is exactly the Riemann integral $\int_a^b f(x)dx$. In Prelims Analysis III, we have used a different approach by using step functions to define upper and lower integrals, and we can show that the two methods lead to the same upper and lower integrals, and thus the Riemann integrals.

**Lemma 10.2** *$f : [a, b] \to \mathbb{R}$ is Riemann integrable on $[a, b]$, if and only if*

$$\lim_{|D| \to 0} \sum_j (M_j - m_j)(x_j - x_{j-1}) = 0$$

*where $M_j = \sup_{x \in [x_{j-1}, x_j]} f(x)$ and $m_j = \inf_{x \in [x_{j-1}, x_j]} f(x)$ for $j = 1, \cdots, n$ for any given finite partition $D = \{a = x_0 < \cdots < x_n = b\}$.*

**Proof.** In fact, if $f$ is Riemann integrable, then $f$ is bounded as proved in the previous lemma. We here provide another proof. By definition, there is $\delta > 0$ such that for any positive integer$N$ such that $\frac{1}{N}(b - a) \leq \delta$ we have

$$\int_a^b f(x)dx - 1 \leq \sum_{j=1}^{N} f(\xi_j) \frac{b - a}{N} \leq \int_a^b f(x)dx + 1 \tag{10.2}$$

for any $\xi_j \in [x_{N,j-1}, x_{N,j}]$, $j = 1, \cdots, N$, where $x_{N,j} = a + \frac{j}{N}(b - a)$ as before. Fix such $N$. Let $C_1 = \max_{j=1,\cdots,N} |f(x_{N,j-1})|$. Then, for any fixed $j_0 = 1, \cdots, N$, applying the previous inequality to $\xi_j = x_{N,j-1}$ for $j \neq j_0$, we have

$$\begin{aligned} f(\xi_{j_0}) \frac{b - a}{N} &\leq \int_a^b f(x)dx + 1 - \sum_{j \neq j_0}^{N} f(x_{N,j-1}) \frac{b - a}{N} \\ &\leq \int_a^b f(x)dx + 1 + \frac{N - 1}{N}(b - a)C_1 \end{aligned}$$

and similarly, we have

$$f(\xi_{j_0})\frac{b-a}{N} \geq \int_a^b f(x)dx - 1 - \frac{N-1}{N}(b-a)C_1.$$

Therefore

$$\frac{N}{b-a}\left(\int_a^b f(x)dx - 1\right) - NC_1 \leq m_{j_0} \leq M_{j_0} \leq \frac{N}{b-a}\left(\int_a^b f(x)dx + 1\right) + NC_1$$

where $m_j$ and $M_j$ are infinimum and supremum of $f$ over $[x_{N,j-1}, x_{N,j}]$. But this is true for any $j_0 = 1, \cdots, N$, so that

$$\frac{N}{b-a}\left(\int_a^b f(x)dx - 1\right) - NC_1 \leq \inf_{[a,b]} f(x) \leq \sup_{[a,b]} f(x) \leq \frac{N}{b-a}\left(\int_a^b f(x)dx + 1\right) + NC_1$$

which provides explicit bounds for $f$ in terms of its integral.

Since $f$ is Riemann integrable on $[a,b]$, for every $\varepsilon > 0$, there is $\delta > 0$ such that for any finite partition $D : a = x_0 < \cdots < x_n = b$ such that $|D| < \delta$ we have

$$\int_a^b f(x)dx - \varepsilon \leq \sum_{j=1}^n f(\xi_j)(x_j - x_{j-1}) \leq \int_a^b f(x)dx + \varepsilon.$$

For any but fixed $\xi_j \in [x_{j-1}, x_j]$ for $j = 2, 3, \cdots, n$, taking supremum and infinimum over $\xi_1 \in [x_0, x_1]$ in the above inequality we obtain

$$\begin{aligned}\int_a^b f(x)dx - \varepsilon &\leq m_1(x_1 - x_0) + \sum_{j=2}^n f(\xi_j)(x_j - x_{j-1}) \\ &\leq M_1(x_1 - x_0) + \sum_{j=2}^n f(\xi_j)(x_j - x_{j-1}) \leq \int_a^b f(x)dx + \varepsilon\end{aligned}$$

and repeat the process $n$ times we obtain that

$$\int_a^b f(x)dx - \varepsilon \leq \sum_{j=1}^n m_j(x_j - x_{j-1}) \leq \sum_{j=1}^n M_j(x_j - x_{j-1}) \leq \int_a^b f(x)dx + \varepsilon. \tag{10.3}$$

In particular

$$\sum_{j=1}^n (M_j - m_j)(x_j - x_{j-1}) \leq 2\varepsilon$$

and therefore, by definition,

$$\lim_{|D|\to 0} \sum_{j=1}^n (M_j - m_j)(x_j - x_{j-1}) = 0.$$

From the previous inequality (10.3), we may also conclude that

$$\lim_{|D|\to 0} \sum_{j=1}^n m_j(x_j - x_{j-1}) = \lim_{|D|\to 0} \sum_{j=1}^n M_j(x_j - x_{j-1}) = \int_a^b f(x)dx.$$

Conversely, if it holds that

$$\lim_{|D|\to 0}\sum_{j=1}^{n}(M_j-m_j)(x_j-x_{j-1})=0,$$

then for every $\varepsilon>0$, there is $\delta>0$ such that

$$0\leq\sum_{j=1}^{n}(M_j-m_j)(x_j-x_{j-1})<\varepsilon$$

for any finite partition $D=\{a=x_0<\cdots<x_n=b\}$ such that $|D|<\delta$. In particular, applying the above to any but fixed partition $D$ with $|D|<\delta$, the inequality above implies that both $m_i$ and $M_i$ are finite, so that $f$ is bounded on $[a,b]$. Moreover

$$0\leq\overline{\int_a^b}f(x)dx-\underline{\int_a^b}f(x)dx\leq\sum_{j=1}^{n}(M_j-m_j)(x_j-x_{j-1})<\varepsilon$$

so that

$$\overline{\int_a^b}f(x)dx-\underline{\int_a^b}f(x)dx=0.$$

By definition, $f$ is Riemann integrable on $[a,b]$. ■

If $J=\langle a,b\rangle$ is a finite interval with ends $a<b$, but not necessary closed one, then naturally, we say $f$ is Riemann integrable on $J$, if $f1_{[a,b]}$ is Riemann integrable, where $g=f1_{[a,b]}$ is the extension of $f$ so that $g(a)$ and / or $g(b)=0$ if $f(a)$ and / or $f(b)$ are not defined. In Prelims Analysis III, we have proved the following: if $f$ is Riemann integrable on $[a,b]$, and $g$ is a function on $[a,b]$ which agrees with $f$ except for finite many points, then $g$ is also Riemann integrable on $[a,b]$ and $\int_a^b f(x)dx=\int_a^b g(x)dx$. It is also proved that a bounded function $f$ on $[a,b]$ which is continuous on $(a,b)$ is Riemann integrable. With the help of Lebesgue's measure, we are now in a position to prove the final result of this kind of statements about Riemannian integrals.

To this end, let us introduce some notations which are used in the proof of our main result about Riemann integrals.

Suppose $f$ is a bounded function on $[a,b]$. For every $n$, we define the dyadic partition $a=x_0^{(n)}<x_1^{(n)}<\cdots<x_{2^n}^{(n)}=b$, where $x_k^{(n)}=a+\frac{k}{2^n}(b-a)$ for $k=0,1,\cdots,2^n$. Define two sequences of step functions $\varphi_n(x)=m_k^{(n)}$ and $\psi_n(x)=M_k^{(n)}$ if $x\in(x_{k-1}^{(n)},x_k^{(n)}]$ and $\varphi_n(a)=\psi_n(a)=f(a)$, where $m_k^{(n)}=\inf\{f(x):x\in[x_{k-1}^{(n)},x_k^{(n)}]\}$ and similarly $M_k^{(n)}=\sup\{f(x):x\in[x_{k-1}^{(n)},x_k^{(n)}]\}$. Then $\varphi_n\uparrow$ and $\psi_n\downarrow$. Moreover $\varphi_n\leq f\leq\psi_n$ on $[a,b]$ for every $n$. Let $f_{\inf}=\lim\varphi_n$ and $f_{\sup}=\lim\psi_n$. Then both $f_{\inf}$ and $f_{\sup}$ are measurable functions on $[a,b]$, and are bounded on $[a,b]$, so they must be Lebesgue integrable. Moreover $f_{\inf}\leq f\leq f_{\sup}$ on $[a,b]$. Suppose $C$ is a bound of $|f|$, then $|\varphi_n|,|\psi_n|\leq C$. Therefore, by Lebesgue's Dominated Convergence theorem, we have

$$\int_{[a,b]}f_{\inf}dm=\lim_{n\to\infty}\int_a^b\varphi_n(x)dx=\underline{\int_a^b}f(x)dx$$

and

$$\int_{[a,b]}f_{\sup}dm=\lim_{n\to\infty}\int_a^b\psi_n(x)dx=\overline{\int_a^b}f(x)dx.$$

It follows that

$$0 \leq \overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx = \int_{[a,b]} (f_{\sup} - f_{\inf})dm.$$

Therefore we have the following

**Lemma 10.3** *Let $f$ be a bounded function on $[a,b]$. Then*

*1) both functions $f_{\sup}$ and $f_{\inf}$ are bounded and measurable on $[a,b]$, and therefore both are Lebesgue integrable on $[a,b]$;*

*2) $f_{\inf} \leq f \leq f_{\sup}$ on $[a,b]$;*

*3) $f$ on $[a,b]$ is Riemann integrable if and only if*

$$\int_{[a,b]} (f_{\sup} - f_{\inf})dm = 0$$

*which is equivalent to that $f_{\sup} = f = f_{\inf}$ almost everywhere on $[a,b]$.*

Let $D = \left\{ x_k^{(n)} : k = 0, \cdots, 2^n, n = 1, 2, \cdots \right\}$ be the subset of all possible dyadic points in $[a,b]$. Then $D$ is countable, so that $m(D) = 0$.

**Lemma 10.4** *Under the above notations and assumptions. If $x \in (a,b) \backslash D$, then $f$ is continuous at $x$ if and only if $f_{\sup}(x) = f_{\inf}(x)$.*

**Proof.** Let $x_0 \in [a,b]$. Then, for every $\varepsilon > 0$, there is $N$ such that

$$\varphi_n(x_0) \leq f_{\inf}(x_0) \leq \varphi_n(x_0) + \varepsilon$$

and

$$\psi_n(x_0) - \varepsilon \leq f_{\sup}(x_0) \leq \psi_n(x_0)$$

so that

$$0 \leq \psi_n(x_0) - \varphi_n(x_0) \leq f_{\sup}(x_0) - f_{\inf}(x_0) + 2\varepsilon.$$

for all $n \geq N$. Suppose in addition that $x_0 \in (a,b) \setminus D$, then, for every $n$, there is a unique integer $k_n \in \{0, \cdots, 2^n - 1\}$ such that $x_0 \in (x_{k_n-1}^{(n)}, x_{k_n}^{(n)})$. Let

$$\delta = \min\{x_0 - x_{k_N-1}^{(N)}, x_{k_N}^{(N)} - x_0\}$$

which is positive (as $x_0 \notin D$). Then, for any $x$ such that $|x - x_0| < \delta$, we have

$$\begin{aligned} |f(x) - f(x_0)| &\leq \sup_{|y-x_0|<\delta} f - \inf_{|y-x_0|<\delta} f \\ &\leq M_{k_N}^{(N)} - m_{k_N}^{(N)} = \psi_N(x_0) - \varphi_N(x_0) \end{aligned}$$

so that

$$|f(x) - f(x_0)| \leq f_{\sup}(x_0) - f_{\inf}(x_0) + 2\varepsilon$$

for any $x$ such that $|x - x_0| < \delta$. Therefore, if $f_{\sup}(x_0) = f_{\inf}(x_0)$, then $f$ is continuous at $x_0$.

On the other hand, we have

$$0 \leq f_{\sup}(x_0) - f_{\inf}(x_0) \leq \psi_n(x_0) - \varphi_n(x_0)$$

for any $n$ and $x_0$. If $f$ is continuous at $x_0$, then for every $\varepsilon > 0$, there is $\delta > 0$ such that $|f(x)-f(x_0)| < \frac{\varepsilon}{2}$ for any $x \in (x_0-\delta, x_0+\delta)$. Thus $|f(x)-f(y)| < \varepsilon$ for any $x, y \in (x_0-\delta, x_0+\delta)$. Choose positive integer $N$ such that $\frac{b-a}{2^N} < \frac{1}{2}\delta$, then, if $x_0 \notin D$, for any $n \geq N$ there is a unique $l_n = 0, \cdots, 2^n - 1$ such that $x_0 \in (x_{l_n-1}^{(n)}, x_{l_n}^{(n)}) \subseteq (x_0 - \delta, x_0 + \delta)$, so that

$$M_{l_n}^{(n)} - m_{l_n}^{(n)} \leq \sup_{x,y\in(x_0-\delta,x_0+\delta)} |f(x) - f(y)| \leq \varepsilon.$$

Therefore, for $n \geq N$

$$\begin{aligned} 0 &\leq f_{\sup}(x_0) - f_{\inf}(x_0) \leq \psi_n(x_0) - \varphi_n(x_0) \\ &= M_{l_n}^{(n)} - m_{l_n}^{(n)} \leq \varepsilon \end{aligned}$$

that is $0 \leq f_{\sup}(x_0) - f_{\inf}(x_0) \leq \varepsilon$ for every $\varepsilon > 0$, thus we must have $f_{\sup}(x_0) = f_{\inf}(x_0)$, which completed the proof. ■

**Theorem 10.5** *Let $f : [a, b] \to \mathbb{R}$ be a function. Then $f$ is Riemann integrable on $[a, b]$, if and only if $f$ is* bounded *and $f$ is* continuous almost everywhere *on $[a, b]$.*

**Proof.** Let $A = \{f_{\sup} \neq f_{\inf}\}$. If $f$ is Riemann integrable on $[a, b]$, then $f$ must be bounded on $[a, b]$, and $f_{\sup} = f_{\inf}$ almost everywhere on $[a, b]$, so that $m(A) = 0$, and therefore $A \cup D$ is null. According to Lemma 10.4, $f$ is continuous on $(a, b) \setminus (A \cup D)$. Therefore $f$ is continuous almost everywhere on $[a, b]$. Conversely, if $f$ is continuous almost everywhere and bounded on $[a, b]$, then $f_{\sup} = f_{\inf}$ almost surely, so that $f$ is Riemann integrable on $[a, b]$ according to Lemma 10.3 ■

**Corollary 10.6** *If $f$ is Riemann integrable on $[a, b]$, then $f = f_{\sup} = f_{\inf}$ almost everywhere on $[a, b]$, so that $f$ must be bounded and measurable on $[a, b]$. Moreover*

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

Let us consider an example, the Dirichlet function $f$ on $[0, 1]$, where $f(x) = 0$ if $x \in (0, 1)$ is irrational , and $f(x) = \frac{1}{p+q}$ if $x = \frac{p}{q}$ in lowest term, where $p \geq 0$ and $q > 0$ are two integers. Then $f$ is continuous at any irrational in $[0, 1]$, thus $f$ is continuous almost everywhere on $[0, 1]$, so that $f$ is Riemann integrable. Moreover, since $f = 0$ a.e., thus its Riemann integral $\int_0^1 f(x)dx = 0$.