MEASURE THEORY NOTES

(2010)

Mr. Andrew L Pinchuck

Department of Mathematics (Pure & Applied)

Rhodes University

Contents

Introduction

In mathematics, more specifically in measure theory, a measure on a set is a systematic way to assign to each suitable subset a number, intuitively interpreted as the size of the subset. In this sense, a measure is a generalization of the concepts of length, area, volume, etc. A particularly important example is the Lebesgue measure on a Euclidean space, which assigns the conventional length, area and volume of Euclidean geometry to suitable subsets of R^n. Most measures met in practice in analysis (and in many cases also in probability theory) are Radon measures. Radon measures have an alternative definition in terms of linear functionals on the locally convex space of continuous functions with compact support. In this course, however, we will mainly concentrate on the Lebesgue measure.

Measure theory was developed in successive stages during the late 19th and early 20th century by Emile Borel, Henri Lebesgue, Johann Radon and Maurice Frchet, among others. The main applications of measures are in the foundations of the Lebesgue integral, in Andrey Kolmogorov's axiomatisation of probability theory and in ergodic theory. In integration theory, specifying a measure allows one to define integrals on spaces more general than subsets of Euclidean space; moreover, the integral with respect to the Lebesgue measure on Euclidean spaces is more general and has a richer theory than its predecessor, the Riemann integral. Probability theory considers measures that assign to the whole set, the size 1, and considers measurable subsets to be events whose probability is given by the measure. Ergodic theory considers measures that are invariant under, or arise naturally from, a dynamical system. The importance of integration and how measure theory puts integration and probability theory on an axiomatic foundation is a principle motivation for the development of this theory. Measure theory is useful in functional analysis in defining L^p function spaces, which cannot otherwise be defined properly.

These notes are intended to be an introduction to measure theory and integration. I make no claims of originality with regards to this material, and I have used a number of different sources as references in the compilation of these notes. Chapter 1 deals with the theory of Riemann integration and highlights some of its shortcomings. In Chapter 2, we define the fundamental concepts of measure theory and present some well-known, significant results. In Chapter 3, we define the Lebesgue integral and consider some results that follow from this and lastly, in Chapter 4, we provide a brief introduction to the study of signed measures.

Throughout the course notes you will find exercises that need to be attempted for complete understanding of the text. Once again, I should mention that I have found wikipedia to be a useful resource and recommend its use to students. All the course information and material can be found on the Rhodes mathematics web site: www.ru.ac.za/mathematics by following the appropriate links.

Chapter 1

Riemann integration

1.1 Introduction

In the branch of mathematics known as real analysis, the Riemann integral, created by Bernhard Riemann (1826-1866), was the first rigorous definition of the integral of a function on an interval. The Riemann integral was clearly an improvement upon the way in which integration was initially conceived by Leibnitz and Newton, i.e., simply as the anti-derivative. While the Riemann integral is one of the easiest integrals to define, and has obvious uses in numerical analysis, it is unsuitable for many theoretical purposes. Some of these technical deficiencies can be remedied by the RiemannStieltjes integral, and most of them disappear in the Lebesgue integral. The idea behind the Riemann integral is to use very simple approximations for the area of S. By taking better and better approximations, we can say that 'in the limit', we get exactly the area of S under the curve.

1.2 Basic definitions and theorems

In this section we discuss the construction of the Riemann integral. This discussion reveals some of the shortcomings of the Riemann integral and we present these as a motivation for the construction of the Lebesgue integral.

1.2.1 Definition

Let $[a, b]$ be a closed interval in $\mathbb{R}$. A **partition of** $[a, b]$ is a set $P = \{x_0, x_1, x_2, \ldots, x_n\}$ of points in $\mathbb{R}$ such that

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b.$$

Let f be a real-valued function which is bounded on $[a, b]$ and let $P = \{x_0, x_1, x_2, \ldots, x_n\}$ be a partition of $[a, b]$. Denote by

$$M = \sup_{a \le x \le b} f(x) \text{ and } m = \inf_{a \le x \le b} f(x).$$

Since f is bounded on $[a, b]$, it is bounded on each subinterval $[x_{i-1}, x_i]$, for each $i = 1, 2, \ldots, n$. Let

$$M_i = \sup\{f(x) : x_{i-1} \le x \le x_i\} \text{ and } m_i = \inf\{f(x) : x_{i-1} \le x \le x_i\},$$

for each $i = 1, 2, \ldots, n$. Clearly,

$$m \le m_i \le M_i \le M, \text{ for each } i = 1, 2, \ldots, n.$$

We can now form the sums

$$U(f, P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) \text{ and } L(f, P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}).$$

1.2.2 Definition
The sums $U(f, P)$ and $L(f, P)$ are called, respectively, the **upper** and the **lower sum** of f relative to the partition P.

1.2.3 Remark

[1] It is important to note that $U(f, P)$ and $L(f, P)$ *depend* on the partition P. If f is nonnegative on $[a, b]$, then the upper sum $U(f, P)$ is the sum of the areas of the rectangles whose heights are M_i and whose bases are $[x_{i-1}, x_i]$. Similarly, $L(f, P)$ is the sum of the areas of the rectangles whose heights are m_i, and whose bases are $[x_{i-1}, x_i]$.

[2] It also follows that $U(f, P) \geq L(f, P)$.

1.2.4 Theorem
Let f be a real-valued function which is bounded on $[a, b]$ and let $P = \{x_0, x_1, x_2, \ldots, x_n\}$ be a partition of $[a, b]$. Then
$$m(b-a) \leq L(f, P) \leq U(f, P) \leq M(b-a).$$

PROOF.

Since $M_i \leq M$, for each $i = 1, 2, \ldots, n$, it follows that

$$U(f, P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) \leq \sum_{i=1}^{n} M(x_i - x_{i-1}) = M \sum_{i=1}^{n}(x_i - x_{i-1}) = M(b-a).$$

Similarly, since $m \leq m_i$, for each $i = 1, 2, \ldots, n$, it follows that

$$L(f, P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}) \geq \sum_{i=1}^{n} m(x_i - x_{i-1}) = m \sum_{i=1}^{n}(x_i - x_{i-1}) = m(b-a).$$

□

This theorem says that the set $A = \{U(f, P) : P \text{ is a partition of } [a, b]\}$ is bounded below by $m(b-a)$. Hence, A has an infimum, $\Sigma(f)$, say. That is,

$$\Sigma(f) = \inf_P U(f, P),$$

where the infimum is taken over all possible partitions P of $[a, b]$. This theorem also shows that the set $B = \{L(f, P) : P \text{ is a partition of } [a, b]\}$ is bounded above by $M(b-a)$ and hence B has a supremum, $\sigma(f)$, say. That is,

$$\sigma(f) = \sup_P L(f, P),$$

where the supremum is taken over all possible partitions P of $[a, b]$. It follows that

$$m(b-a) \leq \Sigma(f) \leq M(b-a), \text{ and}$$
$$m(b-a) \leq \sigma(f) \leq M(b-a).$$

1.2.5 Definition
Let f be a real-valued function which is bounded on $[a, b]$. The **upper integral of** f **on** $[a, b]$ is defined by

$$\overline{\int_a^b} f(x)dx = \inf_P U(f, P),$$

and the **lower integral of f on** $[a,b]$ is defined by

$$\underline{\int_a^b} f(x)dx = \sup_P L(f,P),$$

where, once again, the infimum and supremum are taken over all possible partitions P of $[a,b]$.

It is intuitively clear that $\underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx$. Of course, this is not enough and we will therefore prove this fact shortly.

1.2.6 Definition
Let $P = \{x_0, x_1, x_2, \ldots, x_n\}$ be a partition of $[a,b]$. A partition P^* of $[a,b]$ is called a **refinement** of P, denoted by $P \subseteq P^*$, if $x_i \in P^*$, for each $i = 0, 1, 2, \ldots, n$. A partition P^* is called a **common refinement** of the partitions P_1 and P_2 of $[a,b]$ if P^* is a refinement of both P_1 and P_2.

The next theorem states that refining a partition decreases the upper sum and increases the lower sum.

1.2.7 Lemma (Refinement Lemma)
Let f be a real-valued function which is bounded on $[a,b]$. If P^* is a refinement of a partition $P = \{x_0, x_1, x_2, \ldots, x_n\}$ of $[a,b]$, then

$$L(f,P) \leq L(f,P^*) \leq U(f,P^*) \leq U(f,P).$$

PROOF.

Suppose that P^* has one more point that P, say a point x^* which lies in the subinterval $[x_{r-1}, x_r]$. Let

$$L_1 = \sup\{f(x) : x_{r-1} \leq x \leq x_*\}, L_2 = \sup\{f(x) : x^* \leq x \leq x_r\} \text{ and}$$
$$l_1 = \inf\{f(x) : x_{r-1} \leq x \leq x_*\}, l_2 = \inf\{f(x) : x^* \leq x \leq x_r\}.$$

Recalling that

$$M_r = \sup\{f(x) : x_{r-1} \leq x \leq x_r\}, \text{ and } m_r = \inf\{f(x) : x_{r-1} \leq x \leq x_r\},$$

we observe that

$$m_r \leq l_1, m_r \leq l_2, L_1 \leq M_r, \text{ and } L_2 \leq M_r.$$

It follows that

$$m_r(x_r - x_{r-1}) = m_r(x_r - x^*) + m_r(x^* - x_{r-1}) \leq l_2(x_r - x^*) + l_1(x^* - x_{r-1}).$$

Hence,

$$\begin{aligned}
L(f,P^*) &= \sum_{j=1}^{r-1} m_j(x_j - x_{j-1}) + l_1(x^* - x_{r-1}) + l_2(x_r - x^*) + \sum_{j=r+1}^{n} m_j(x_j - x_{j-1}) \\
&\geq \sum_{j=1}^{r-1} m_j(x_j - x_{j-1}) + m_r(x_r - x_{r-1}) + \sum_{j=r+1}^{n} m_j(x_j - x_{j-1}) \\
&= \sum_{j=1}^{n} m_j(x_j - x_{j-1}) = L(f,P).
\end{aligned}$$

Similarly,

$$M_r(x_r - x_{r-1}) = M_r(x_r - x^*) + M_r(x^* - x_{r-1}) \geq L_2(x_r - x^*) + L_1(x^* - x_{r-1}),$$

and so

$$\begin{aligned}
U(f, P) &= \sum_{j=1}^{r-1} M_j(x_j - x_{j-1}) + L_1(x^* - x_{r-1}) + L_2(x_r - x^*) + \sum_{j=r+1}^{n} M_j(x_j - x_{j-1}) \\
&\leq \sum_{j=1}^{r-1} M_j(x_j - x_{j-1}) + M_r(x_r - x_{r-1}) + \sum_{j=r+1}^{n} M_j(x_j - x_{j-1}) \\
&= \sum_{j=1}^{n} M_j(x_j - x_{j-1}) = U(f, P).
\end{aligned}$$

the case where P^* contains $k \geq 2$ more points than P can be proved by simply repeating the above argument k times.

□

1.2.8 Theorem

Let f be a real-valued function which is bounded on $[a, b]$. Then,

$$\underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx.$$

PROOF.

Let P_1 and P_2 be any two partitions of $[a, b]$ and let P^* be their common refinement. Then by Lemma 1.2.7,

$$L(f, P_1) \leq L(f, P^*) \leq U(f, P^*) \leq U(f, P_2).$$

Since P_1 is any partition of $[a, b]$, it follows that

$$\sup_P L(f, P) \leq U(f, P_2),$$

and since P_2 is any partition of $[a, b]$, we have that

$$\sup_P L(f, P) \leq \inf_P U(f, P),$$

where the infimum and the supremum are taken over all possible partitions P of $[a, b]$. Thus,

$$\underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx.$$

□

1.2.9 Remark

Implicit in the proof of Theorem 1.2.8, is the fact that no lower sum can exceed an upper sum. That is, every lower sum is less than or equal to every upper sum.

1.2.10 Definition

Let f be a real-valued function on $[a, b]$. We say that f is **Riemann-integrable** on $[a, b]$ if f is bounded on $[a, b]$ and

$$\underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx.$$

If f is Riemann-integrable on $[a,b]$, we define **the integral of** f **on** $[a,b]$ to be the common value of the upper and lower integrals, i.e,

$$\int_a^b f(x)dx = \underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx.$$

We shall denote by $\mathcal{R}[a,b]$ the collection of all functions that are Riemann-integrable on $[a,b]$.

1.2.11 Remark

In the definition of the integral of f on $[a,b]$, we have tacitly assumed that $a < b$. If $a = b$, then $\int_a^b f(x)dx = \int_a^a f(x)dx = 0$. Also if $b < a$, then define

$$\int_b^a f(x)dx = -\int_a^b f(x)dx.$$

1.2.12 Examples

[1] Show that if f is a constant function on $[a,b]$, then $f \in \mathcal{R}[a,b]$ and find it's integral.

Solution: Let $f(x) = k$, for all $x \in [a,b]$ and let $P = \{x_0, x_1, x_2, \ldots, x_n\}$ by any partition of $[a,b]$. Then

$$M_i = \sup\{f(x) : x_{i-1} \le x \le x_i\} = k \text{ and } m_i = \inf\{f(x) : x_{i-1} \le x \le x_i\} = k,$$

for each $i = 1, 2, \ldots, n$. Therefore,

$$U(f,P) = \sum_{i=1}^n M_i(x_i - x_{i-1}) = k\sum_{i=1}^n (x_i - x_{i-1}) = k(b-a) \text{ and}$$
$$L(f,P) = \sum_{i=1}^n m_i(x_i - x_{i-1}) = k\sum_{i=1}^n (x_i - x_{i-1}) = k(b-a)$$

Since P is *any* partition of $[a,b]$, it follws that $U(f,P) = L(f,P) = k(b-a)$, for all partitions P of $[a,b]$. Therefore,

$$\underline{\int_a^b} f(x)dx = k(b-a) = \overline{\int_a^b} f(x)dx.$$

That is, f is integrable on $[a,b]$ and

$$\int_a^b f(x)dx = k(b-a).$$

[2] Let f be a function defined by

[3]

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \cap [0,1] \\ 0 & \text{if } x \notin \mathbb{Q} \cap [0,1]. \end{cases}$$

Show that f is not Riemann-integrable on $[0,1]$.

Solution: Firstly we observe that f is bounded on $[0,1]$. Let $P = \{x_0, x_1, x_2, \dots, x_n\}$ be any partition of $[0,1]$. Since for each $i = 1, 2, \dots, n$, the subinterval $[x_{i-1}, x_i]$ contains both rational and irrational numbers, we have that

$$\begin{aligned} M_i &= \sup\{f(x) : x_{i-1} \le x \le x_i\} = \sup\{0,1\} = 1, \text{ and} \\ m_i &= \inf\{f(x) : x_{i-1} \le x \le x_i\} = \inf\{0,1\} = 0, \end{aligned}$$

for each $i = 1, 2, \dots, n$. Thus,

$$\begin{aligned} U(f,P) &= \sum_{i=1}^{n} M_i(x_i - x_{i-1}) = (1)\sum_{i=1}^{n}(x_i - x_{i-1}) = 1 - 0 = 1, \text{ and} \\ L(f,P) &= \sum_{i=1}^{n} m_i(x_i - x_{i-1}) = (0)\sum_{i=1}^{n}(x_i - x_{i-1}) = 0. \end{aligned}$$

Since P is *any* partition of $[0,1]$, it follows that $U(f,P) = 1 - 0 = 1$ and $L(f,P) = 0$, for all partitions of P of $[0,1]$. Therefore,

$$\underline{\int_0^1} f(x)dx = 0 \text{ and } \overline{\int_0^1} f(x)dx = 1,$$

and hence f is not Riemann-integrable on $[0,1]$.

Notice that in the last example, the given function f is discontinuous at every point.

1.2.13 Theorem (Darboux's integrability condition)
Let f be a real-valued function which is bounded on $[a,b]$. Then f is integrable on $[a,b]$ if and only if, for any $\epsilon > 0$, there exists a partition P^* of $[a,b]$ such that

$$U(f,P^*) - L(f,P^*) < \epsilon.$$

PROOF.

Assume that f is integrable on $[a,b]$ and let $\epsilon > 0$. Since

$$\int_a^b f(x)dx = \underline{\int_a^b} f(x)dx = \sup_P L(f,P),$$

there is a partition P_1 of $[a,b]$ such that

$$\int_a^b f(x)dx - \frac{\epsilon}{2} < L(f,P_1).$$

Again, since

$$\int_a^b f(x)dx = \overline{\int_a^b} f(x)dx = \inf_P U(f,P),$$

there exists a partition P_2 of $[a,b]$ such that

$$U(f,P_2) < \int_a^b f(x)dx + \frac{\epsilon}{2}.$$

Let P^* be a common refinement of P_1 and P_2. Then

$$\int_a^b f(x)dx - \frac{\epsilon}{2} < L(f, P_1) \le L(f, P^*) \le U(f, P^*) \le U(f, P_2) < \int_a^b f(x)dx + \frac{\epsilon}{2}.$$

It now follows that

$$U(f, P^*) - L(f, P^*) < \epsilon.$$

For the converse, assume that given $\epsilon > 0$, there is a partition P^* of $[a, b]$ such that

$$U(f, P^*) - L(f, P^*) < \epsilon.$$

Then,

$$\overline{\int_a^b} f(x)dx = \inf_P U(f, P) \le U(f, P^*), \text{ and } \underline{\int_a^b} f(x)dx = \sup_P L(f, P) \ge L(f, P^*).$$

Thus,

$$0 \le \overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx \le U(f, P^*) - L(f, P^*) < \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have that

$$\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx.$$

That is, $f \in \mathcal{R}[a, b]$.

□

We now highlight the following important fact which is contained in the first part of the proof of Darboux's integrability condition:

1.2.14 Theorem

If f is integrable on $[a, b]$, then for each $\epsilon > 0$, there exists a partition P of $[a, b]$ such that

$$\int_a^b f(x)dx - \epsilon < L(f, P) \le U(f, P) < \int_a^b f(x)dx + \epsilon.$$

1.2.15 Theorem

If f is continuous on $[a, b]$, then it is integrable there.

PROOF.

Since f is continuous on $[a, b]$, we have that f is bounded on $[a, b]$. Furthermore, f is uniformly continuous on $[a, b]$. Hence, given $\epsilon > 0$, there is a $\delta > 0$ such that

$$|f(x) - f(y)| < \frac{\epsilon}{b - a}, \text{ whenever } x, y \in [a, b] \text{ and } |x - y| < \delta.$$

Let $P = \{x_0, x_1, x_2, \ldots, x_n\}$ be any partition of $[a, b]$ such that $x_i - x_{i-1} < \delta$, for each $i = 1, 2, \ldots, n$. By the Extreme-Value Theorem (applied to f on $[x_{i-1}, x_i]$, for each $i = 1, 2, \ldots, n$), there exist points t_i and s_i in $[x_{i-1}, x_i]$, for each $i = 1, 2, \ldots, n$ such that

$$f(t_i) = \sup\{f(x) : x_{i-1} \le x \le x_i\} = M_i \text{ and } f(s_i) = \inf\{f(x) : x_{i-1} \le x \le x_i\} = m_i.$$

Since $x_i - x_{i-1} < \delta$, it follows that $|t_i - s_i| < \delta$, and so

$$M_i - m_i = f(t_i) - f(s_i) = |f(t_i) - f(s_i)| < \frac{\epsilon}{b-a}, \text{ for all } i = 1, 2, \ldots, n.$$

Thus,

$$\begin{aligned} U(f,P) - L(f,P) &= \sum_{i=1}^{n} M_i(x_i - x_{i-1}) - \sum_{i=1}^{n} m_i(x_i - x_{i-1}) = \sum_{i=1}^{n}(M_i - m_i)(x_i - x_{i-1}) \\ &\le \sum_{i=1}^{n} |f(t_i) - f(s_i)|(x_i - x_{i-1}) < \sum_{i=1}^{n}\left(\frac{\epsilon}{b-a}\right)(x_i - x_{i-1}) \\ &= \frac{\epsilon}{b-a}(b-a) = \epsilon. \end{aligned}$$

It now follows from Theorem 1.2.13, that f is integrable on $[a,b]$.

□

1.2.16 Theorem
If f is monotone on $[a,b]$, then f is integrable there.

PROOF.

Assume that f is monotone increasing on $[a,b]$ and $f(a) < f(b)$. Since $f(a) \le f(x) \le f(b)$, for all $x \in [a,b]$, f is clearly bounded on $[a,b]$. We want to show that, given any $\epsilon > 0$, there is a partition P of $[a,b]$ such that $U(f,P) - L(f,P) < \epsilon$. Let $\epsilon > 0$ be given an let $P = \{x_0, x_1, x_2, \ldots, x_n\}$ be any partition of $[a,b]$ such that $x_i - x_{i-1} < \frac{\epsilon}{f(b)-f(a)}$, for each $i = 1, 2, \ldots, n$. Since f is increasing on $[a,b]$, we have that

$$M_i = \sup\{f(x) : x_{i-1} \le x \le x_i\} = f(x_i) \text{ and } m_i = \inf\{f(x) : x_{i-1} \le x \le x_i\} = f(x_{i-1}),$$

for each $i = 1, 2, \ldots, n$. Hence,

$$\begin{aligned} U(f,P) - L(f,P) &= \sum_{i=1}^{n} M_i(x_i - x_{i-1}) - \sum_{i=1}^{n} m_i(x_i - x_{i-1}) = \sum_{i=1}^{n}(M_i - m_i)(x_i - x_{i-1}) \\ &\le \sum_{i=1}^{n}[f(x_i) - f(x_{i-1})](x_i - x_{i-1}) < \frac{\epsilon}{f(b) - f(a)} \sum_{i=1}^{n}[f(x_i) - f(x_{i-1})] \\ &= \frac{\epsilon}{f(b) - f(a)}[f(b) - f(a)] = \epsilon. \end{aligned}$$

It then follows from Theorem 1.2.13, that f is Riemann-integrable on $[a,b]$. The case where f is monotone decreasing can be proved in exactly the same way.

□

1.2.17 Theorem
If f is integrable on $[a,b]$ and $a \le c < d \le b$, then f is integrable on $[c,d]$.

PROOF.

Since f is integrable on $[a,b]$, it is bounded there. Hence f is bounded on $[c,d]$. Furthermore, given any $\epsilon > 0$, there is a partition P of $[a,b]$, such that

$$U(f,P) - L(f,P) < \epsilon.$$

Let $P^* = P \cup \{c,d\}$. Then P^* is a refinement of P, and hence

$$U(f,P^*) - L(f,P^*) \le U(f,P) - L(f,P) < \epsilon.$$

Let $Q_1 = P^* \cap [a, c]$, $Q_2 = P^* \cap [c, d]$, $Q_3 = P^* \cap [d, b]$. Then $P^* = Q_1 \cup Q_2 \cup Q_3$, and so,

$$U(f, P^*) = U(f, Q_1) + U(f, Q_2) + U(f, Q_3), \text{ and}$$
$$L(f, P^*) = L(f, Q_1) + L(f, Q_2) + L(f, Q_3).$$

Hence,

$$[U(f, Q_1) - L(f, Q_1)] + [U(f, Q_2) - L(f, Q_2)] + [U(f, Q_3) - L(f, Q_3)] = U(f, P^*) - L(f, P^*) < \epsilon.$$

Note that all terms on the left are nonnegative. Therefore Q_2 is a partition of $[c, d]$ with the property that

$$U(f, Q_2) - L(f, Q_2) < \epsilon.$$

This implies that f is integrable on $[c, d]$.

□

1.2.18 Corollary
If f is integrable on $[a, b]$ and $a < c < b$, then f is integrable on both $[a, c]$ and $[c, b]$.

The following theorem states that the converse of Corollary 1.2.18 also holds.

1.2.19 Theorem
If $a < c < b$ and f is integrable on both $[a, c]$ and $[c, b]$, then f is integrable on $[a, b]$ and

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

PROOF.

Let $\epsilon > 0$ be given. Since f is integrable on $[a, c]$ and on $[c, b]$, there are partitions P_1 and P_2 of $[a, c]$ and $[c, b]$ respectively, such that

$$U(f, P_1) - L(f, P_1) < \frac{\epsilon}{2}, \text{ and}$$
$$U(f, P_2) - L(f, P_2) < \frac{\epsilon}{2}.$$

Let $P = P_1 \cup P_2$. Then P is a partition of $[a, b]$ and

$$\begin{aligned} U(f, P) - L(f, P) &= U(f, P_1) + U(f, P_2) - L(f, P_1) - L(f, P_2) \\ &= U(f, P_1) - L(f, P_1) + U(f, P_2) - L(f, P_2) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Therefore f is integrable on $[a, b]$. Furthermore,

$$\begin{aligned} \int_a^b f(x)dx \le U(f, P) &= U(f, P_1) + U(f, P_2) \\ &< \left[L(f, P_1) + \frac{\epsilon}{2}\right] + \left[L(f, P_2) + \frac{\epsilon}{2}\right] \\ &= L(f, P_1) + L(f, P_2) + \epsilon \\ &\le \int_a^c f(x)dx + \int_c^b f(x)dx + \epsilon. \end{aligned}$$

Since ϵ is arbitrary, it follows that

$$\int_a^b f(x)dx \leq \int_a^c f(x)dx + \int_c^b f(x)dx \tag{1.1}$$

Also,

$$\begin{aligned}\int_a^b f(x)dx \geq L(f,P) &= L(f,P_1) + L(f,P_2) \\ &> \left[U(f,P_1) - \frac{\epsilon}{2}\right] + \left[U(f,P_2) - \frac{\epsilon}{2}\right] \\ &= U(f,P_1) + U(f,P_2) - \epsilon\end{aligned}$$

Since ϵ is arbitrary, it follows that

$$\int_a^b f(x)dx \geq \int_a^c f(x)dx + \int_c^b f(x)dx. \tag{1.2}$$

Combining (1.1) and (1.2), we have that

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

□

1.2.20 Corollary
Let f be defined on $[a,b]$ and suppose that $a = c_0 < c_1 < \cdots < c_{n-1} < c_n = b$. Then f is integrable on $[a,b]$ if and only if f is integrable on $[c_{k-1}, c_k]$, for each $k = 1, 2, \ldots, n$. In this case,

$$\int_a^b f(x)dx = \sum_{k=1}^{n} \int_{c_{k-1}}^{c_k} f(x)dx.$$

1.2.21 Corollary
If f is continuous at all but a finite set of points in $[a,b]$, then f is Riemann-integrable on $[a,b]$.

1.3 Fundamental theorem of calculus

1.3.1 Theorem (Fundamental theorem of calculus)
Let f be a bounded real-valued integrable function on $[a,b]$. Suppose that $F : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b) and that

$$F'(x) = f(x), \text{ for each } x \in (a,b).$$

Then,

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

PROOF.

Let $P = \{x_0, x_1, x_2, \ldots, x_n\}$ be a partition of $[a, b]$.

$$M_i = \sup\{f(x) : x_{i-1} \le x \le x_i\} \text{ and } m_i = \inf\{f(x) : x_{i-1} \le x \le x_i\},$$

for each $i = 1, 2, \ldots, n$. For each $i = 1, 2, \ldots, n$, F is continuous on $[x_{i-1}, x_i]$ and differentiable on (x_{i-1}, x_i). By the Mean Value Theorem, there is a $\zeta_i \in (x_{i-1}, x_i)$ such that

$$F(x_i) - F(x_{i-1}) = F'(\zeta_i)(x_{i-1} - x_i) = f(\zeta_i)(x_{i-1} - x_i).$$

Since, for each $i = 1, 2, \ldots, n$ $m_i \le f(\zeta_i) \le M_i$, it follows that

$$\begin{aligned} m_i(x_i - x_{i-1}) \le f(\zeta_i)(x_{i-1} - x_i) \le M_i(x_i - x_{i-1}) &\Leftrightarrow m_i(x_i - x_{i-1}) \le F(x_i) - F(x_{i-1}) \le M_i(x_i - x_{i-1}) \\ &\Rightarrow \sum_{i=1}^{n} m_i(x_i - x_{i-1}) \le \sum_{i=1}^{n}[F(x_i) - F(x_{i-1})] \le \sum_{i=1}^{n} M_i(x_i - x_{i-1}) \\ &\Leftrightarrow L(f, P) \le F(b) - F(a) \le U(f, P). \end{aligned}$$

Since P is an *arbitrary* partition of $[a, b]$, it follows that $F(b) - F(a)$ is an upper bound for the set $\{L(f, P) : P \text{ is a partition of } [a, b]\}$ and a lower bound for the set $\{U(f, P) : P \text{ is a partition of } [a, b]\}$. Therefore,

$$\underline{\int_a^b} f(x)dx = \sup_P L(f, P) \le F(b) - F(a) \le \inf_P U(f, P) = \overline{\int_a^b} f(x)dx.$$

Since f is integrable on $[a, b]$, we have that

$$\underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx = \int_a^b f(x)dx,$$

and consequently $\displaystyle\int_a^b f(x)dx = F(b) - F(a)$

□

1.3.2 Exercise

[1] Let f be the function on $[0, 1]$ given by

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

Show that f is not Riemann-integrable on $[0, 1]$.

[2] Let f be the function on $[0, 1]$ given by

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x < \frac{1}{2} \\ x - \frac{1}{2} & \text{if } \frac{1}{2} \le x < 1. \end{cases}$$

(a) Show, from first principles, that f is Riemann-integrable on $[0, 1]$.

(b) Quote a result that assures us that f is Riemann-integrable.

(c) Find $\int_0^1 f(x)dx$.

(d) Let $\{r_1, r_2, r_3, \dots\}$ be an enumeration of rationals in the interval $[0, 1]$. For each $n \in \mathbb{N}$, define

$$f_n(x) = \begin{cases} 1 & \text{if } x \in \{r_1, r_2, r_3, \dots, r_n\} \\ 0 & \text{otherwise.} \end{cases}$$

Show that (f_n) is a nondecreasing sequence of functions that are Riemann-integrable on $[0, 1]$. Show also that the sequence (f_n) converges pointwise to the function

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \cap [0, 1] \\ 0 & \text{otherwise,} \end{cases}$$

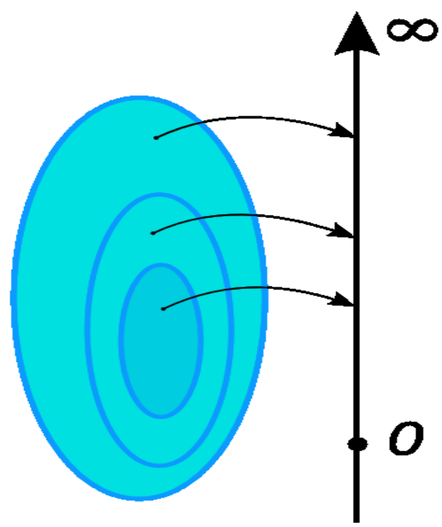
and that f is not Riemann-integrable.

Chapter 2

Measure spaces

2.1 Introduction

We now concern ourselves with the idea of a defining a measure on the real numbers, which puts the intuitive idea of length on an axiomatic foundation. This turns out to be more difficult than you might expect and we begin with the motivation for developing measure theory. If we are to properly define the definite integral of a real-valued function, we need a meaningful notion of the length of a set $A \subset \mathbb{R}$.



If we consider the set of real numbers, we would like to be able to define a function that measures the 'length' of a set. That is a function $m : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ such that:

[1] $\mu([a, b]) = b - a$, for $a \leq b$.

[2] μ is translation invariant, i.e., $\mu(A) = m(x + A)$, for $A \subset \mathbb{R}$ and $x \in \mathbb{R}$.

[3] $\mu(\emptyset) = 0$.

[4] If (A_i) is a sequence of mutually disjoint sets in $(\mathbb{R})$ (i.e., $A_i \cap A_j = \emptyset$, for every $i \neq j$), then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} A_i.$$

This turns out to be impossible. That is, you cannot define such a function whose domain consists of the whole of $\mathcal{P}(\mathbb{R})$. To show this we present the following example by Giuseppe Vitali (1905):

2.1.1 Example

A Vitali set V, is a subset of $[0, 1]$, which for each real number r, V contains exactly one number v such that $v - r$ is rational (This implies that V is uncountable, and also, that $v - u$ is irrational, for any $u, v \in V$, $u \neq v$). Such sets can be shown to exist provided we assume the axiom of choice.

To construct the Vitali set V, consider the additive quotient group $\mathbb{R}/\mathbb{Q}$. Each element of this group is a 'shifted copy' of the rational numbers. That is, a set of the form $\mathbb{Q} + r$, for some $r \in \mathbb{R}$. Therefore, the elements of this group are subsets of $\mathbb{R}$ and partition $\mathbb{R}$. There are uncountably many elements. Since each element intersects $[0, 1]$, we can use the axiom of choice to choose a set $V \subset [0, 1]$ containing exactly one representative out of each element $\mathbb{R}/\mathbb{Q}$.

A Vitali set is non-measurable. To show this, we argue by contradiction and assume that V is measurable. Let $q_1, q_2, \ldots$ be an enumeration of the rational numbers in $[-1, 1]$. From the construction of V, note that the translated sets $V_k = V + q_k = \{v + q_k : v \in V\}$, $k = 1, 2, \ldots$ are pairwise disjoint, and further note that $[0, 1] \subseteq \bigcup_k V_k \subseteq [-1, 2]$ (to see the first inclusion, consider any real number r in $[0, 1]$ and let v be the representative in V for the equivalence class $[r]$, the $r - v = q$ for some rational number q in $[-1, 1]$).

Now assume that there exists a μ that satisfies the four conditions listed above. Apply μ to these inclusions using condition [4]:

$$1 \leq \sum_{k=1}^{\infty} \mu(V_k) \leq 3.$$

Since μ is translation invariant, $\mu(V_k) = \mu(V)$ and

$$1 \leq \sum_{k=1}^{\infty} \mu(V) \leq 3.$$

This is impossible. Summing infinitely many copies of the constant $\mu(V)$ yields either zero or infinity, according to whether the constant is zero or positive. In neither case is the sum in $[1, 3]$. So, V cannot have been measurable after all, i.e., μ must not define any value for $\lambda(V)$.

It is important to stress that this whole argument depends on the assumption of the axiom of choice.

2.2 Notation and preliminaries

Let X be a set.

- We will use the standard notation for the *compliment of A*:

$$A^c = \{x \in X : x \notin A\}$$

- We denote the *complement of B with respect to A* by $A \backslash B = A \cap B^c$.
- We denote by $\mathcal{P}(X)$, the *powerset* of X, i.e., the set of all subsets of X.
- Let A and B be in $\mathcal{P}(X)$. The *symmetric difference* of A and B, denoted by $A \triangle B$, is given by

$$A \triangle B = (A \backslash B) \cup (B \backslash A).$$

- A subset $\mathcal{C}$ of $\mathcal{P}(X)$ is said to be:

[1] *closed under finite intersections* if, whenever $\{A_1, A_2, \ldots, A_n\}$ is a finite collection of elements of $\mathcal{C}$, then $\displaystyle\bigcap_{j=1}^{n} A_j \in \mathcal{C}$.

[2] *closed under finite unions* if, whenever $\{A_1, A_2, \ldots, A_n\}$ is a finite collection of elements of $\mathcal{C}$, then $\displaystyle\bigcup_{j=1}^{n} A_j \in \mathcal{C}$.

[3] *closed under countable intersections* if, whenever (A_n) is a sequence of elements of $\mathcal{C}$, then $\displaystyle\bigcap_{j=1}^{\infty} A_j \in \mathcal{C}$.

[4] *closed under countable unions* if, whenever (A_n) is a sequence of elements of $\mathcal{C}$, then $\displaystyle\bigcup_{j=1}^{\infty} A_j \in \mathcal{C}$.

- Let (A_n) be a sequence in $\mathcal{P}(X)$. We denote by

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m, \text{ and } \liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m.$$

2.2.1 Exercise

Let X be a nonempty set. Prove that if a collection $\mathcal{C}$ of $\mathcal{P}(X)$ is closed under countable intersections (resp., unions), then it is closed under finite intersections (resp., unions).

2.2.2 Definition

An **algebra** (in X) is a collection Σ of subsets of X such that:

(a) $\emptyset, X \in \Sigma$.

(b) If $A \in \Sigma$, then $A^c \in \Sigma$.

(c) If $A_1, \ldots, A_n \in \Sigma$, then $\cup_{i=1}^{n} A_i$ and $\cap_{i=1}^{n} A_i$ are in Σ.

Σ is a σ**-algebra** (or σ**-field**) if in addition:

(d) If $A_1, A_2, \ldots$ are in Σ, then $\cup_{i=1}^{\infty} A_i$ and $\cap_{i=1}^{\infty} A_i$ are in Σ.

2.2.3 Examples

[1] Let X be any nonempty set. Then there are two special σ-algebras on X: $\Sigma_1 = \{\emptyset, X\}$ and $\Sigma_2 = \mathcal{P}(X)$ (the *power set of* X).

[2] Let $X = [0,1]$. Then $\Sigma = \{\emptyset, X, [0, \frac{1}{2}], (\frac{1}{2}, 1]\}$ is a σ-algebra on X.

[3] Let $X = \mathbb{R}$ and let

$$\Sigma = \{A \subset \mathbb{R} : A \text{ is countable or } A^c \text{ is countable}\}.$$

Σ is a σ-algebra on X. Parts (a) and (b) of the definition are easy to show. Suppose $A_1, A_2, \ldots$ are all in Σ. If each of the A_i's are countable, then $\cup_i A_i$ is countable,and so in Σ. If $A_{i_0}^c$ is countable for some i_0, then

$$(\cup_i A_i)^c = \cap_i A_i^c \subset A_{i_0}^c$$

is countable, and again $\cup_i A_i$ is in Σ. Since $\cap_i A_i = (\cup_i A_i^c)^c$, then the countable intersection of sets in Σ is again in Σ.

2.2.4 Remark

Let $X = \mathbb{R}$ and let Σ be a σ-algebra on X. If (A_n) is a sequence of sets in Σ, then $\liminf_{n\to\infty} A_n \in \Sigma$.

If $\mathcal{S}$ is any collection of subsets of a set X, then by taking the intersection of all σ-algebras containing $\mathcal{S}$, we obtain the smallest σ-algebra containing $\mathcal{S}$. Such a σ-algebra, denoted by $\sigma(\mathcal{S})$ is said to be **generated by S**. It is clear that if $\mathcal{S}$ is a σ-algebra, then $\sigma(\mathcal{S}) = \mathcal{S}$.

2.2.5 Definition
The smallest σ-algebra conataining all open subsets of $\mathbb{R}$ is called the **Borel** σ-algebra of $\mathbb{R}$ and is denoted by $\mathcal{B}(\mathbb{R})$. Elements of this σ-algebra are called **Borel sets**.

2.2.6 Proposition
Let $\mathcal{I} = \{[a,b) : a, b \in \mathbb{R}, a \leq b\}$. Then $\sigma(\mathcal{I}) = \mathcal{B}(\mathbb{R})$.

PROOF.

If $a, b \in \mathbb{R}$ and $a \leq b$, then since

$$[a,b) = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b\right),$$

it follows that $[a,b) \in \mathcal{B}(\mathbb{R})$ and consequently, $\mathcal{I} \subseteq \mathcal{B}(\mathbb{R})$. It now follows that $\sigma(\mathcal{I}) \subseteq \mathcal{B}(\mathbb{R})$.

Conversely, let A be an open subset of $\mathbb{R}$. Then A is a countable union of open intervals. Since any open interval (a,b) can be expressed in the form

$$(a,b) = \bigcup_{n=n_0}^{\infty} \left[a + \frac{1}{n}, b\right),$$

where $\frac{1}{n_0} < b - a$, it follows that $A \in \sigma(\mathcal{I})$. Thus $\sigma(\mathcal{I}) = \mathcal{B}(\mathbb{R})$.

□

2.2.7 Exercise
Let X be a nonempty set and $\mathcal{S}$ the smallest σ-algebra in X which contains all singletons $\{x\}$. Show that

$$\mathcal{S} = \{A \in \mathcal{P}(X) : \text{ either } A \text{ or } A^c \text{ is countable}\}.$$

2.3 Measure on a set

2.3.1 Definition
A **measurable space** is a pair (X, Σ), where X is a set and Σ is a σ-algebra on X. The elements of the σ-algebra Σ are called **measurable sets**.

2.3.2 Definition
Let (X, Σ) be a measurable space. A **measure** on (X, Σ) is a function $\mu : \Sigma \to [0, \infty]$ such that:

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

(b) if (A_i) is a sequence of mutually disjoint sets in Σ (i.e., $A_i \in \Sigma$, for each $i = 1, 2, 3, \ldots$, and $A_i \cap A_j = \emptyset$ for every $i \neq j$), then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} A_i.$$

Property (b) in the above definition is known as the *countable additivity* property of the measure μ. If (b) is replaced by the condition that $\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} A_i$ for every finite sequence $A_1, A_2, \ldots, A_n$ of pairwise disjoint sets in Σ, then we say that μ is *finitely additive*. It follows immediately, that if a measure is countably additive, then it must also be finitely additive. If you have a finite sequence $A_1, A_2, \ldots, A_n$ in Σ, then you can consider it to be an infinite sequence (A_i) with $A_{n+1} = A_{n+2} = \ldots \emptyset$. Then

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

2.3.3 Definition

A **measure space** is a triple (X, Σ, μ), where X is a set, Σ a σ-algebra in X, and μ a measure on Σ.

2.3.4 Remark

Let (X, Σ, μ) be a measure space. If $\mu(X) = 1$, then (X, Σ, μ) is called a **probability space** and μ is called a **probability measure**.

2.3.5 Examples

[1] Let X be a nonempty set and $\Sigma = \mathcal{P}(X)$. Define $\nu : \Sigma \to [0, \infty]$ by

$$\nu(A) = \begin{cases} n & \text{if } A \text{ has } n \text{ elements} \\ \infty & \text{otherwise.} \end{cases}$$

Then ν is a measure on X called the **counting measure**.

[2] Let (X, Σ) be a measurable space. Choose and fix $x \in X$. Define $\mu_x : \Sigma \to [0, \infty]$ by

$$\mu_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise.} \end{cases}$$

Then μ_x is a measure on X called the **unit point mass at** x.

2.3.6 Theorem

Let (X, Σ, μ) be a measure space. Then the following statements hold:

[1] If $A, B \in \Sigma$ and $A \subseteq B$, then $\mu(A) \leq \mu(B)$.

[2] If (A_n) is a sequence in Σ, then

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

[3] If (A_n) is a sequence in Σ such that $A_n \subseteq A_{n+1}$, for each $n \in \mathbb{N}$, then

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

[4] If (A_n) is a sequence in Σ such that $A_n \supseteq A_{n+1}$, for each $n \in \mathbb{N}$ and $\mu(A_1) < \infty$, then

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

PROOF.

[1] Write $B = A \cup (B \backslash A)$, a disjoint union. Since μ is finitely additive, we have that

$$\mu(B) = \mu(A) + \mu(B \backslash A) \geq \mu(A).$$

[2] Set $B_1 = A_1$, and for each natural number $n > 1$, let $B_n = A_n \backslash \bigcup_{i=1}^{n-1} A_i$. Then

$$B_n = \bigcap_{i=1}^{n-1} (A_n \backslash A_i) = A_n \bigcap_{i=1}^{n-1} A_i^c.$$

Therefore, for each $n = 1, 2, 3, \ldots$, we have that $B_n \in \Sigma$. Clearly, $B_n \subseteq A_n$, for each $n = 1, 2, 3, \ldots$. Let l and k be integers such that $l < k$. Then $B_l \subset A_l$, and so

$$\begin{aligned} B_l \cap B_k \subset A_l \cap B_k &= A_l \cap A_k \bigcap_{i=1}^{k-1} A_i^c \\ &= A_l \cap A_k \cap \cdots \cap A_l^c \cap \cdots \\ &= A_l \cap A_l^c \cap A_k \cap \cdots \cap \cdots = \emptyset \cap A_k \cap \cdots \cap \cdots \\ &= \emptyset. \end{aligned}$$

That is, $B_i \bigcap_{i \neq j} B_j = \emptyset$.

Claim: $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$.

Since $B_n \subseteq A_n$, for each $n = 1, 2, 3, \ldots$, we have that $\bigcup_{n=1}^{\infty} B_n \subseteq \bigcup_{n=1}^{\infty} A_n$.

Let $x \in \bigcup_{n=1}^{\infty} A_n$. Then $x \in A_n$ for some n. Let k be the smallest index such that $x \in A_n$. Then $x \in A_k$ and $x \notin A_j$, for each $j = 1, 2, \ldots, k-1$. Therefore, $x \in A_j^c$, for each $j = 1, 2, \ldots, k-1$. That is, $x \in B_k = A_k \bigcap_{j=1}^{k-1} A_j^c$, whence $x \in \bigcup_{n=1}^{\infty} B_n$, which proves the claim.

Therefore,

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

by part [1].

[3] Let $B_1 = A_1$ and $B_n = A_n \backslash A_{n-1}$ for all $n > 1$. Then $B_i \bigcap_{i \neq j} B_j = \emptyset, A_k = \bigcup_{n=1}^{k} B_n$ and $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$. Therefore,

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_{k\to\infty} \sum_{n=1}^{k} \mu(B_n) = \lim_{k\to\infty} \mu(A_k).$$

[4] Let $A = \bigcap_{n=1}^{\infty} A_n$. Then

$$A_1 \backslash A = A_1 \backslash \bigcap_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} (A_1 \backslash A_n),$$

and $A_1 \backslash A_n \subset A_1 \backslash A_{n+1}$, for each natural number n. Therefore, by [3] above,

$$\mu(A_1 \backslash A) = \mu\left(\bigcup_{n=1}^{\infty} (A_1 \backslash A_n)\right) = \lim_{n\to\infty} \mu(A_1 \backslash A_n). \tag{2.1}$$

Since $A \subset A_1$ and $\mu(A_1) < \infty$, we have that $\mu(A) < \infty$. Also, since $A_n \subset A_1$, for each natural number n, it follows that $\mu(A) < \infty$, for each n. Writing A_1 as a disjoint union $A_1 = (A_1 \backslash A) \cup A$, we have that $\mu(A_1) = \mu(A_1 \backslash A) + \mu(A)$. It now follows that

$$\mu(A_1 \backslash A) = \mu(A_1) - \mu(A). \tag{2.2}$$

Similarly, by expressing A_1 as $A_1 = (A_1 \backslash A_n) \cup A_n$, we get that

$$\mu(A_1 \backslash A_n) = \mu(A_1) - \mu(A_n). \tag{2.3}$$

From equations 2.1, 2.2 and 2.3, we have that

$$\mu(A_1) - \mu(A) = \lim_{n\to\infty} [\mu(A_1) - \mu(A_n)] = \mu(A_1) - \lim_{n\to\infty} \mu(A_n).$$

Since $\mu(A_1) < \infty$, we have that

$$\mu(A) = \lim_{n\to\infty} \mu(A_n), \text{ that is } \mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

□

2.3.7 Theorem (Borel-Cantelli Lemma)
Let (X, Σ, μ) be a measure space. If (A_n) is a sequence of measurable sets such that $\sum_{n=1}^{\infty} \mu(A_n) < \infty$, then $\mu\left(\limsup_{n\to\infty} A_n\right) = 0$.
PROOF.

Since $\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$, it follows that for each $n \in \mathbb{N}$,

$$\mu\left(\limsup_{n\to\infty} A_n\right) \le \mu\left(\bigcup_{n=1}^{\infty} A_k\right) \le \sum_{n=1}^{\infty} \mu(A_k) \to 0$$

as $n \to \infty$.

□

2.3.8 Definition
Let (X, Σ, μ) be a measure space. The measure μ is said to be:

[1] **finite** if $\mu(X) < \infty$, and

[2] σ**-finite** if there is a sequence (X_n) of sets in Σ with $\mu(X_n) < \infty$, for each n such that $X = \bigcup_{n=1}^{\infty} X_n$.

2.4 Outer measure

Constructing a measure is a nontrivial exercise as illustrated in the introduction to this chapter. The basic idea that we will use is as follows: We will begin by defining an 'outer measure' on all subsets of a set. Then we will restrict the outer measure to a reasonable class of subsets and in so doing we will define a measure. This is indeed how Lebesgue solved the problems associated with measure and integration. Both were published as part of his dissertation in 1902.

2.4.1 Definition
Let X be a set. An **outer measure** on X is a set function $\mu^* : \mathcal{P}(X) \to [0.\infty]$ satisfying the following properties:

(a) $\mu^*(\emptyset) = 0$.

(b) If $A \subseteq B$, then $\mu^*(A) \le \mu^*(B)$.

(c) For any sequence (A_n) of subsets of X, $\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \le \sum_{n=1}^{\infty} \mu^*(A_n)$.

Property (b) above is referred to as *monotonicity*, and property (c) is *countable subadditivity*

2.4.2 Examples

[1] Let X be a set. Define μ^* on $\mathcal{P}(X)$ by:

$$\mu^*(A) = \begin{cases} 1 & \text{if } A \neq \emptyset \\ 0 & \text{if } A = \emptyset. \end{cases}$$

Then μ^* is an outer measure on X.

[2] Let X be an uncountable set. Define μ^* on $\mathcal{P}(X)$ by:

$$\mu^*(A) = \begin{cases} 0 & \text{if } A \text{ is countable} \\ 1 & \text{if } A \text{ is uncountable.} \end{cases}$$

Then μ^* is an outer measure on X.

We can now use the concept of outer measure to construct a measure.

2.4.3 Definition

Let X be a set and μ^* an outer measure on X. A subset E of X is said to be **μ^*-measurable** if, for every subset A of X, we have

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

2.4.4 Remark

[1] Since $A = (A \cap E) \cup (A \cap E^c)$ and μ^* is subadditive, we always have that

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

Therefore, a subset E of X is μ^*-measurable if, for every subset A of X,

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

[2] Replacing E by E^c in the definition of μ^*-measurability, we have that

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

Thus, E is μ^*-measurable if and only if E^c is μ^*-measurable.

2.4.5 Lemma

Let X be a set, μ^* an outer measure on X and E a subset of X. If $\mu^*(E) = 0$, then E is μ^*-measurable.

PROOF.

Let A be any subset of X. Then $A \cap E \subseteq E$. Since μ^* is monotone, $\mu^*(A \cap E) \leq \mu^*(E) = 0$. Also, since $A \cap E^c \subseteq A$, it follows that $\mu^*(A \cap E^c) \leq \mu^*(A)$. Therefore,

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \cap E^c),$$

and therefore, E is μ^*-measurable.

$\square$

2.4.6 Corollary

The empty set $\emptyset$ is μ^*-measurable.

2.4.7 Lemma

Let X be a set, μ^* an outer measure on X, and E and F μ^*-measurable subsets of X. Then $E \cup F$ is μ^*-measurable.

PROOF.

Let A be any susbset of X. Then since E is μ^*-measurable,

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

Also, since F is μ^*-measurable,

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

From the two equations above we get:

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

Therefore,

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

whence $E \cup F$ is μ^* measurable.

□

2.4.8 Theorem
Let X be a set and μ^* an outer measure on X. Denote by $\mathcal{M}$, the collection of all μ^*-measurable subsets of X. Then $\mathcal{M}$ is an algebra on X.

PROOF.

This follows from Remark 2.4.4 [2], Corollary 2.4.6 and Lemma 2.4.7

□

2.4.9 Proposition
Let X be a set and μ^* an outer measure on X. If $E_1, E_2 \in \mathcal{M}$ and $E_1 \cap E_2 = \emptyset$, then

$$\mu^*[A \cap (E_1 \cup E_2)] = \mu^*(A \cap E_1) + \mu^*(A \cap E_2).$$

PROOF.

Since E_1 is μ^*-measurable,

$$\mu^*(A) = \mu^*(A \cap E_1) + \mu^*(A \cap E_1^c),$$

for each subset A of X. Replacing A by $A \cap (E_1 \cup E_2)$, we have

$$\begin{aligned}\mu^*[A \cap (E_1 \cup E_2)] &= \mu^*[(A \cap (E_1 \cup E_2)) \cap E_1] + \mu^*[(A \cap (E_1 \cup E_2)) \cap E_1^c]\\ &= \mu^*[((A \cap E_1) \cup (A \cap E_2)) \cap E_1] + \mu^*[((A \cap E_1) \cup (A \cap E_2)) \cap E_1^c]\\ &= \mu^*[(A \cap E_1 \cap E_1) \cup (A \cap E_2 \cap E_1)] + \mu^*[(A \cap E_1 \cap E_1^c) \cup (A \cap E_2 \cap E_1^c)]\\ &= \mu^*(A \cap E_1) + \mu^*(A \cap E_2).\end{aligned}$$

□

2.4.10 Theorem
Let X be a set and μ^* an outer measure on X. Suppose that (E_n) be a sequence of mutually disjoint sets in $\mathcal{M}$. Then for any $n \geq 1$ and any subset A of X,

$$\mu^*\left(A \cap \bigcup_{j=1}^{n} E_j\right) = \sum_{j=1}^{n} \mu^*(A \cap E_j). \tag{2.4}$$

PROOF.

We prove this theorem by induction. The case $n = 1$ is obviously true. The case $n = 2$ has been proved in Proposition 2.4.9. Assume that

$$\mu^*\left(A \cap \bigcup_{j=1}^{n-1} E_j\right) = \sum_{j=1}^{n-1} \mu^*(A \cap E_j). \tag{2.5}$$

Firstly, we note that

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

Since E_n is μ^*-measurable,

$$\begin{aligned}
\mu^*\left[A \cap \left(\bigcup_{j=1}^{n} E_j\right)\right] &= \mu^*\left[A \cap \left(\bigcup_{j=1}^{n} E_j\right) \cap E_n\right] + \mu^*\left[A \cap \left(\bigcup_{j=1}^{n} E_j\right) \cap E_n^c\right] \\
&= \mu^*(A \cap E_n) + \mu^*\left[A \cap \left(\bigcup_{j=1}^{n-1} E_j\right)\right] \\
&= \mu^*(A \cap E_n) + \sum_{j=1}^{n-1} \mu^*(A \cap E_j) \text{ (by 2.5)} \\
&= \sum_{j=1}^{n} \mu^*(A \cap E_j).
\end{aligned}$$

□

By taking $A = X$ in (2.4), we obtain

$$\mu^*\left(\bigcup_{j=1}^{n} E_j\right) = \sum_{j=1}^{n} \mu^*(E_j).$$

i.e., μ^* is finitely additive on $\mathcal{M}$.

The following theorem is the key to using an outer measure to construct a measure.

2.4.11 Theorem

Let X be a set and μ^* an outer measure on X. Denote by $\mathcal{M}$, the collection of all μ^*-measurable subsets of X. Then

[1] $\mathcal{M}$ is a σ-algebra, and

[2] the restriction of $\mu^*|_{\mathcal{M}}$ of μ^* to $\mathcal{M}$ is a measure on $(X, \mathcal{M})$.

PROOF.

[1] We have already shown in Theorem 2.4.8 that $\mathcal{M}$ is an algebra. It remains to show that $\mathcal{M}$ is closed under countable unions. We will show that $\mathcal{M}$ is in fact closed under *pairwise disjoint* countable

unions. Let (E_n) be a pairwise disjoint sequence of sets in $\mathcal{M}$, $S_n = \bigcup_{j=1}^{n} E_j$ and $S = \cap_{j=1}^{\infty} E_j$. Then, since $S_n \in \mathcal{M}$, we have that for any $A \subseteq X$,

$$\begin{aligned}
\mu^*(A) &= \mu^*(A \cap S_n) + \mu^*(A \cap S_n^c) \\
&\geq \mu^*(A \cap S_n) + \mu^*(A \cap S^c) \quad \text{(since } S_n^c \supset S^c \text{ and } \mu^* \text{ is monotone)} \\
&= \sum_{j=1}^{n} \mu^*(A \cap E_j) + \mu^*(A \cap S^c) \quad \text{(by Theorem 2.4.10).}
\end{aligned}$$

Since this is true for every natural number n, it follows that

$$\begin{aligned}
\mu^*(A) &\geq \sum_{j=1}^{\infty} \mu^*(A \cap E_j) + \mu^*(A \cap S^c) \\
&\geq \mu^*(A \cap S) + \mu^*(A \cap S^c) \quad \text{(since } \mu^* \text{ is countably subadditive).}
\end{aligned}$$

Therefore, $\bigcup_{j=1}^{\infty} E_j \in \mathcal{M}$.

[2] By definition of the outer measure, $\mu(\emptyset) = 0$. We have to show that μ^* is countably additive. We do this as follows, let (E_n) be a pairwise disjoint sequence of sets in $\mathcal{M}$. As shown in [1], for any $A \subseteq X$,

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

If A is replaced by $\cup_{j=1}^{\infty} E_j$, we have that

$$\mu^*\left(\bigcup_{j=1}^{\infty} E_j\right) \geq \sum_{j=1}^{\infty} \mu^*(E_j). \tag{2.6}$$

Since μ^* is, by definition, countably subadditive, we have that

$$\mu^*\left(\bigcup_{j=1}^{\infty} E_j\right) \leq \sum_{j=1}^{\infty} \mu^*(E_j). \tag{2.7}$$

From (2.6) and (2.7), we conclude that

$$\mu^*\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu^*(E_j).$$

Thus, μ^* is a measure on $\mathcal{M}$.

□

The measure $\mu = \mu^*|_{\mathcal{M}}$ on X is often referred to as *the measure on X induced by the outer measure* μ^*.

2.4.12 Exercise

[1] In proving Theorem 2.4.11 [1], we showed that $\mathcal{M}$ is closed under *pairwise disjoint* countable unions. Deduce from this that $\mathcal{M}$ is closed under countable unions.

[2] Let μ^* be an outer measure on a set X such that $\mu^*(A) = 0$, for some subset A of X. Show that if B is a subset of X, then $\mu^*(A \cup B) = \mu^*(B)$.

2.5 Lebesgue measure on $\mathbb{R}$

We now use the approach discussed in the previous section to construct the Lebesgue measure on $\mathbb{R}$. The Lebesgue measure will lead to the definition of the Lebesgue integral in the next chapter. Let $\mathcal{I}$ be a collection of open intervals in $\mathbb{R}$. For $I \in \mathcal{I}$, let $l(I)$ denote the length of I.

2.5.1 Theorem
For any subset A of $\mathbb{R}$, define $m^* : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ by

$$m^*(A) = \inf\left\{ \sum_n l(I_n) : I_n \in \mathcal{I}, \text{for each } n \text{ and } A \subseteq \bigcup_n I_n \right\}.$$

Then m^* is an outer measure on $\mathbb{R}$.

PROOF.

(i) Since $\emptyset \subseteq (a, a)$, for any real number a, it follows that $0 \leq m^*(\emptyset) \leq a - a = 0$. Thus, $m^*(\emptyset) = 0$.

(ii) Let A and B be subsets of $\mathbb{R}$ such that $A \subseteq B$. If $B \subseteq \bigcup_n I_n$, where (I_n) is a sequence of open intervals in $\mathbb{R}$, then $A \subseteq \bigcup_n I_n$. Therefore, $m^*(A) \leq \sum_{n=1}^{\infty} l(I_n)$. It follows that $m^*(A) \leq m^*(B)$.

(iii) Let (A_n) be a sequence of subsets of $\mathbb{R}$. If $m^*(A_n) = \infty$ for some n, then

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

Suppose that $m^*(A_n) < \infty$, for each n. Then given $\epsilon > 0$, there is a sequence $(I_{nk})_k$ of open intervals in $\mathbb{R}$ such that $A_n \subseteq \bigcup_{k=1}^{\infty} I_{nk}$, and

$$\sum_{k=1}^{\infty} l(I_{nk}) < m^*(A_n) + \frac{\epsilon}{2^n}.$$

Now, since $\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} I_{nk}$, it follows that

$$m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} l(I_{nk}) < \sum_{n=1}^{\infty}\left(m^*(A_n) + \frac{\epsilon}{2^n}\right) = \sum_{n=1}^{\infty} m^*(A_n) + \epsilon.$$

Since ϵ is arbitrary, we have that $m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} m^*(A_n)$.

□

The outer measure defined in Theorem 2.5.1 is called the **Lebesgue outer measure**.

2.5.2 Definition
A set $E \subseteq \mathbb{R}$ is said to be **Lebesgue-measurable** (or simply **measurable**) if

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

for every set $A \subseteq \mathbb{R}$.

Denote by $\mathcal{M}$, the set of all Lebesgue-measurable subsets of $\mathbb{R}$. We have already shown in Theorem 2.4.11, that $\mathcal{M}$ is a σ-algebra and that $m = m^*|_{\mathcal{M}}$ is a measure on $\mathcal{M}$. This measure is called the **Lebesgue measure** on $\mathbb{R}$.

The Lebesgue outer measure has other important properties that are listed in the following propositions.

2.5.3 Proposition
The outer measure of an interval I is its length. That is, $m^*(I) = l(I)$.

PROOF.

Case 1: Let I be a closed and bounded interval, say $I = [a, b]$. Then $[a, b] \subset (a - \epsilon, b + \epsilon)$, for each $\epsilon > 0$. Thus,

$$m^*([a, b]) \le l(a - \epsilon, b + \epsilon) = b - a + 2\epsilon.$$

Since ϵ is arbitrary, $m^*(I) = m^*([a, b]) \le b - a$. It remains to show that $m^*(I) \ge b - a$. Suppose that $I \subset \bigcup_{k=1}^{\infty} I_k$, where I_k is open for each $k \in \mathbb{N}$. Since I is compact, there is a finite subcollection $\{I_k\}_{k=1}^m$ of intervals such that $I \subset \bigcup_{k=1}^m$. Since $a \in I$, there is an interval (a_1, b_1) in the collection $\{I_k\}_{k=1}^m$ such that $a_1 < a < b_1$. If $b \le b_1$, then we are done. Otherwise, $a < b_1 < b$ and so $b_1 \in I$. Therefore, there is an interval (a_2, b_2) in the collection $\{I_k\}_{k=1}^m$ such that $a_2 < b_1 < b_2$. If $b < b_2$, then we are done. Otherweise, $a < b_2 < b$. That is, $b_2 \in I$. So there is an interval (a_3, b_3) in the collection $\{I_k\}_{k=1}^m$ such that $a_3 < b_2 < b_3$. Continuing in this fashion, we obtain a sequence of intervals $(a_1, b_1), (a_2, b_2), \ldots, (a_r, b_r)$ from the collection $\{I_k\}_{k=1}^m$ such that $a_i < b_{i-1} < b_i$ for $i = 2, 3, \ldots, r$. Since $\{I_k\}_{k=1}^m$ is a finite collection, this process must terminate after $s \le m$ steps with $b < b_s$.

Now,

$$\begin{aligned}
\sum_{j=1}^{\infty} l(I_j) \ge \sum_{j=1}^{m} l(I_j) \ge \sum_{j=1}^{s} l(a_j, b_j)& \\
&= (b_s - a_s) + (b_{s-1} - a_{s-1}) + (b_{s-2} - a_{s-2}) + \cdots + (b_1 - a_1) \\
&= b_s - (a_s - b_{s-1}) - (a_{s-1} - b_{s-2}) - \cdots - (a_2 - b_1) - a_1 \\
&> b_s - a_1 > b - a.
\end{aligned}$$

It follows from this that $m^*(I) \ge b - a = l(I)$.

Case 2: Let I be an interval of the form (a, b), $(a, b]$ or $[a, b)$. Then for each $0 < \epsilon < (b - a)$ with

$$[a + \frac{\epsilon}{2}, b - \frac{\epsilon}{2}] \subset I \subset [a - \frac{\epsilon}{2}, b + \frac{\epsilon}{2}].$$

Hence,

$$\begin{aligned}
&m^*[a + \frac{\epsilon}{2}, b - \frac{\epsilon}{2}] \le m^*(I) \le m^*[a - \frac{\epsilon}{2}, b + \frac{\epsilon}{2}] \\
\Leftrightarrow\ & l\big([a + \frac{\epsilon}{2}, b - \frac{\epsilon}{2}]\big) \le m^*(I) \le \big([a - \frac{\epsilon}{2}, b + \frac{\epsilon}{2}]\big) \\
\Leftrightarrow\ & b - a - \epsilon \le m^*(I) \le b - a + \epsilon \qquad \text{(by case 1).}
\end{aligned}$$

Since ϵ is arbitrary, it follows that $m^*(I) = b - a = l(I)$.

Case 3: Suppose that I is an unbounded interval. Then for any real number k, there is a closed interval $J \subset I$ such that $l(J) = k$. Hence,

$$m^*(I) \ge m^*(J) = l(J) = k.$$

Thus, $m^*(I) = \infty$.

□

2.5.4 Proposition
For each $x \in \mathbb{R}$, $m^*(\{x\}) = 0$.

PROOF.

Given any $\epsilon > 0$, $\{x\} \subset (x - \frac{\epsilon}{2}, x + \frac{\epsilon}{2})$. Thus

$$0 \geq m^*(\{x\}) \leq m^*(x - \frac{\epsilon}{2}, x + \frac{\epsilon}{2}) = l(x - \frac{\epsilon}{2}, x + \frac{\epsilon}{2}) = \epsilon.$$

Since ϵ is arbitrary, we have that $m^*(\{x\}) = 0$.

□

2.5.5 Corollary
Every countable subset of $\mathbb{R}$ has outer measure zero.

PROOF.

Let A be a countable subset of $\mathbb{R}$. Then $A = \bigcup_n \{a_n\}$. Therefore,

$$0 \leq m^*(A) = m^*\left(\bigcup_n \{a_n\}\right) \leq \sum_n m^*(\{a_n\}) = 0.$$

Thus, $m^*(A) = 0$.

□

2.5.6 Corollary
The interval $[0, 1]$ is uncountable.

2.5.7 Proposition
The Lebesgue outer measure is translation-invariant. That is, $m^*(A + x) = m^*(A)$ for all $A \subset \mathbb{R}$ and $x \in \mathbb{R}$.

PROOF.

Let (I_n) be a sequence of open intervals such that $A \subset \bigcup_n I_n$. Then $A + x \subset \bigcup_n (I_n + x)$. Thus,

$$m^*(A + x) \leq \sum_n l(I_n + x) = \sum_n l(I_n).$$

That is, $m^*(A + x)$ is a lower bound for the sums $\sum_n l(I_n)$, with $A \subset \bigcup_n I_n$. Therefore,

$$m^*(A + x) \leq m^*(A). \tag{2.8}$$

To prove the reverse inequality, we only need to notice that

$$m^*(A) = m^*(A + x + (-x)) \leq m^*(A + x),$$

where the inequality follows from (4.1). Thus, $m^*(A + x) = m^*(A)$.

□

2.5.8 Proposition
For any set $A \subset \mathbb{R}$ and any $\epsilon > 0$, there is an open set V such that $A \subset V$ and $m^*(V) < m^*(A) + \epsilon$.

PROOF.

By definition of $m^*(A)$, there is a sequence (I_n) of open intervals such that $A \subset \bigcup_n I_n$ and

$$\sum_n l(I_n) < m^*(A) + \frac{\epsilon}{2}.$$

If $I_n = (a_n, b_n)$, let $J_n = (a_n - \frac{\epsilon}{2^{n+1}}, b_n)$. Then $A \subset \bigcup_n J_n$. Let $V = \bigcup_n J_n$. Then V is an open set in $\mathbb{R}$ and

$$m^*(V) \leq \sum_n l(J_n) = \sum_n l(I_n) + \frac{\epsilon}{2} < m^*(A) + \epsilon.$$

□

2.5.9 Proposition
For any $a \in \mathbb{R}$, the set (a, ∞) is measurable.

PROOF.

Let A be any subset of $\mathbb{R}$. We must show that

$$m^*(A) \geq m^*(A \cap (a,\infty)) + m^*(A \cap (a,\infty)^c) = m^*(A \cap (a,\infty)) + m^*(A \cap (-\infty, a]).$$

Let $A_1 = A \cap (a,\infty)$ and $A_2 = A \cap (-\infty, a]$. Then we need to show that

$$m^*(A) \geq m^*(A_1) + m^*(A_2).$$

If $m^*(A) = \infty$, then there is nothing to prove. Assume that $m^*(A) < \infty$. From the definition of m^*, given $\epsilon > 0$, there is a sequence (I_n) of open intervals such that $A \subset \bigcup_n I_n$ and

$$m^*(A) + \epsilon > \sum_n l(I_n).$$

Let $J_n = I_n \cap (a,\infty)$ and $J_n' = I_n \cap (-\infty, a]$. Then J_n and J_n' are disjoint (possibly empty) intervals and $I_n = J_n \cup J_n'$. Therefore,

$$l(I_n) = l(J_n) + l(J_n') = m^*(J_n) + m^*(J_n').$$

Since $A_1 \subset \bigcup_n J_n$ and $A_2 \subset \bigcup_n J_n'$, it follows that

$$m^*(A_1) \leq m^*\left(\bigcup_n J_n\right) \leq \sum_n m^*(J_n) = \sum_n l(J_n), \quad \text{and}$$
$$m^*(A_2) \leq m^*\left(\bigcup_n J_n'\right) \leq \sum_n m^*(J_n') = \sum_n l(J_n').$$

Therefore,

$$m^*(A_1) + m^*(A_2) \leq \sum_n [l(J_n) + l(J_n')] = \sum_n l(I_n) < m^*(A) + \epsilon.$$

Since ϵ is arbitrary, $m^*(A_1) + m^*(A_2) \leq m^*(A)$.

□

2.5.10 Corollary

[1] Every interval in $\mathbb{R}$ is measurable.

[2] Every open set in $\mathbb{R}$ is measurable.

[3] Every closed set in $\mathbb{R}$ is measurable.

[4] Every set that is a countable intersection of open sets in $\mathbb{R}$ is measurable.

[5] Every set that is a countable union of closed sets in $\mathbb{R}$ is measurable.

PROOF.

[1] We have already shown that for each $a \in \mathbb{R}$, the interval (a,∞) is measurable. It follows from the fact that $\mathcal{M}$ is an algebra that $(-\infty, a] = (a,\infty)^c$ is measurable for each $a \in \mathbb{R}$.

Since $[a,\infty) = \{a\} \cup (a,\infty)$ and $m^*(\{a\}) = 0$, it follows from Lemma 2.4.5 that $\{a\}$ is measurable. Therefore, $[a,\infty)$ is measurable. Since $\mathcal{M}$ is an algebra, we have that $(-\infty, a) = [a,\infty)^c$ is measurable for each $a \in \mathbb{R}$.

If a and b are real numbers such that $a < b$, then $(a,b) = (-\infty, b) \cap (a,\infty)$. Since $(-\infty, b)$ and (a,∞) are both measurable, it follows thta (a,b) is also measurable.

If a and b are real numbers such that $a < b$, then $[a,b) = \{a\} \cup (a,b)$, $(a,b] = (a,b) \cup \{b\}$, and $[a,b] = \{a\} \cup (a,b) \cup \{b\}$. Therefore, the sets $[a,b)$, $(a,b]$, and $[a,b]$ are all measurable.

[2] Every open set in $\mathbb{R}$ is a countable union of open intervals. Since each open interval is measurable and $\mathcal{M}$ is a σ-algebra, it follows that the union of open intervals is also measurable.

[3] A closed set F in $\mathbb{R}$ is a complement of an open set V in $\mathbb{R}$. That is $F = V^c$ for some open set V in $\mathbb{R}$. Since V is measurable, $V^c = F$ is also measurable.

[4] Let $U = \bigcap_{i=1}^{\infty} V_i$, where V_i is open in $\mathbb{R}$ for each $i \in \mathbb{N}$. Then by [2], V_i is measurable, for each $i \in \mathbb{N}$, and so is V_i^c for each $i \in \mathbb{N}$. Since $\mathcal{M}$ is a σ-algebra, we have that

$$U = \bigcap_{i=1}^{\infty} V_i = \left(\bigcup_{i=1}^{\infty} V_i^c\right)^c$$

is measurable.

[5] Let $F = \bigcup_{i=1}^{\infty} F_i$, where F_i is closed in $\mathbb{R}$, for each $i \in \mathbb{N}$. Then F_i^c is open in $\mathbb{R}$, for each $i \in \mathbb{N}$. By [4], we have that $\bigcap_{i=1}^{\infty} F_i^c$ is measurable. Since $\mathcal{M}$ is a σ-algebra, it follows that $\left(\bigcap_{i=1}^{\infty} F_i^c\right)^c = \bigcup_{i=1}^{\infty} F_i = F$ is measurable.

□

2.5.11 Definition

(a) A set that is a union of a countable collection of closed sets is called a F_σ**-set**.

(b) A set that is an intersection of a countable collection of open sets is called a G_δ**-set**.

2.5.12 Remark

[1] The previous corollary says, amongst other things, that every F_σ-set and every G_δ-set is measurable.

[2] The previous corollary also states that all 'nice' sets in $\mathbb{R}$ are measurable. There are however sets that are not measurable. The point being that these sets are certainly not 'natural' but somewhat strange.

[3] The previous corollary shows that every open set in $\mathbb{R}$ is measurable. The smallest σ-algebra containing all the open sets is called the Borel σ-algebra.

The following result states that any measurable subset of $\mathbb{R}$ can be approximated by the 'nice' ones.

2.5.13 Proposition

Let $E \subset \mathbb{R}$ and m^* the Lebesgue outer measure on $\mathbb{R}$. The following statements are equivalent:

[1] E is measurable.

[2] For each $\epsilon > 0$, there is an open set V such that $E \subset V$ and $m^*(V \backslash E) < \epsilon$.

[3] There is a G_δ-set G such that $E \subset G$ and $m^*(G \backslash E) = 0$.

[4] For each $\epsilon > 0$, there is a closed set F such that $F \subset E$ and $m^*(E \backslash F) < \epsilon$.

[5] There is an F_σ-set H such that $H \subset E$ and $m^*(E \backslash H) = 0$.

PROOF.

[1] ⇒ [2]: Suppose that E is measurable and let $\epsilon > 0$. If $m(E) < \infty$, then by Proposition 2.5.8, there is an open set V such that $E \subset V$ and

$$m^*(V) < m^*(E) + \epsilon.$$

Since V and E are measurable, we can rewrite the equation above as

$$m(V) < m(E) + \epsilon.$$

Hence, $m(V \backslash E) = m(V) - m(E) < \epsilon$. Suppose that $m(E) = \infty$. Write $\mathbb{R} = \bigcup_{n\in\mathbb{N}} I_n$, a union of disjoint finite intervals. For each $n \in \mathbb{N}$, let $E_n = E \cap I_n$. Then $E = \bigcup_{n\in\mathbb{N}} E_n$, and for each $n \in \mathbb{N}$, $m(E_n) < \infty$, Therefore, by the first part, there is an open set V_n such that $E_n \subset V_n$ and $m(V_n \backslash E_n) < \frac{\epsilon}{2^n}$. Write $V = \bigcup_{n\in\mathbb{N}} V_n$. Then V is an open set and

$$V \backslash E = \bigcup_{n\in\mathbb{N}} V_n \backslash \bigcup_{n\in\mathbb{N}} E_n = \bigcup_{n\in\mathbb{N}} (V_n \backslash E_n).$$

Hence

$$m(V \backslash E) \le \sum_{n\in\mathbb{N}} m(V_n \backslash E_n) < \sum_{n\in\mathbb{N}} \frac{\epsilon}{2^n} = \epsilon.$$

[2] ⇒ [3]: For each $n \in \mathbb{N}$, let V_n be an open set such that $E \subset V_n$ and $m^*(V_n \backslash E) < \frac{1}{n}$. Let $G = \bigcap_{n\in\mathbb{N}} V_n$. Then G is a G_δ-set, $E \subset G$, and for each $n \in \mathbb{N}$,

$$m^*(G \backslash E) = m^*\left(\bigcap_{n\in\mathbb{N}} (V_n \backslash E) \right) \le m^*(V_n \backslash E) \le \frac{1}{n}.$$

Thus, $m^*(G \backslash E) = 0$.

[3] ⇒ [1]: Since G is a G_δ-set, it is measurable. Also $G \backslash E$ is measurable since $m^*(G \backslash E) = 0$. Therefore, the set $E = G \backslash (G \backslash E)$ is also measurable.

[1] ⇒ [4]: Suppose that E is measurable and let $\epsilon > 0$. Then E^c is also measurable. Since [1] implies [2], we have, by [2], that there is an open set V such that $E^c \subset V$ and $m^*(V \backslash E^c) < \epsilon$. Note that $V \backslash E^c = E \cap V = E \backslash V^c$. Take $F = V^c$. Then F is a closed set, $F \subset E$, and $m^*(E \backslash F) < \epsilon$.

[4] ⇒ [5]: For each $n \in \mathbb{N}$, let F_n be a closed set such that $F_n \subset E$ and $m^*(E \backslash F_n) < \frac{1}{n}$. Let $H = \bigcup_{n\in\mathbb{N}} F_n$. Then H is an F_σ-set, $H \subset E$, and for each $n \in \mathbb{N}$,

$$m^*(E \backslash H) = m^*\left(E \backslash \bigcup_{n\in\mathbb{N}} F_n \right) = \bigcap_{n\in\mathbb{N}} (E \backslash F_n) \le m^*(E \backslash F_n) < \frac{1}{n}.$$

Thus, $m^*(E \backslash H) = 0$.

[5] ⇒ [1]: Since H is an F_σ-set, it is measurable. Also $E \backslash H$ is measurable since $m^*(E \backslash H) = 0$. Therefore, the set $E = (E \backslash H) \cup H$ is also measurable.

□

Chapter 3

The Lebesgue integral

3.1 Introduction

We now turn our attention to the construction of the Lebesgue integral of general functions, which, as already discussed, is necessary to avoid the technical deficiencies associated with the Riemann integral.

3.2 Measurable functions

3.2.1 Basic notions

The **extended real number system**, $\overline{\mathbb{R}}$, is the set of real numbers together with two symbols $-\infty$ and $+\infty$. That is, $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\} = [-\infty, \infty]$. The algebraic operations for these two infinities are:

[1] For $r \in \mathbb{R}$, $\pm\infty + r = \pm\infty$.

[2] For $r \in \mathbb{R}$, $r(\pm\infty) = \pm\infty$, if $r > 0$ and $r(\pm\infty) = \mp\infty$ if $r < 0$.

[3] $+\infty + (+\infty) = +\infty$ and $-\infty + (-\infty) = -\infty$.

[4] $\infty + (-\infty)$ is undefined.

[5] $0 \cdot (\pm\infty) = 0$.

3.2.1 Definition
Let (X, Σ) be a measurable space and $E \in \Sigma$. A function $f : E \to \overline{\mathbb{R}}$ is said to be **measurable** if for each $\alpha \in \mathbb{R}$, the set $\{x \in E : f(x) > \alpha\}$ is measurable.

3.2.2 Proposition
Let (X, Σ) be a measurable space and $E \in \Sigma$, and $f \to \overline{R}$. Then the following are equivalent:

[1] f is measurable.

[2] For each $\alpha \in \mathbb{R}$, the set $\{x \in E : f(x) \geq \alpha\}$ is measurable.

[3] For each $\alpha \in \mathbb{R}$, the set $\{x \in E : f(x) < \alpha\}$ is measurable.

[4] For each $\alpha \in \mathbb{R}$, the set $\{x \in E : f(x) \leq \alpha\}$ is measurable.

PROOF.

[1] ⇒ [2]: If f is measurable, then the set

$$\{x \in E : f(x) \geq \alpha\} = \bigcap_{n \in \mathbb{N}} \left\{x \in E : f(x) > \alpha - \frac{1}{n}\right\},$$

which is an intersection of measurable sets, is measurable.

[2] ⇒ [3]: If the set $\{x \in E : f(x) \geq \alpha\}$ is measurable, then so is the set

$$E \backslash \{x \in E : f(x) \geq \alpha\} = \{x \in E : f(x) < \alpha\}.$$

[3] ⇒ [4]: If the set $\{x \in E : f(x) < \alpha\}$ is measurable, then so is the set

$$\{x \in E : f(x) \leq \alpha\} = \bigcap_{n \in \mathbb{N}} \left\{x \in E : f(x) < \alpha + \frac{1}{n}\right\}$$

since it is an intersection of measurable sets.

[4] ⇒ [1]: If the set $\{x \in E : f(x) \leq \alpha\}$, is measurable, so is its complement in E. Hence,

$$\{x \in E : f(x) > \alpha\} = E \backslash \{x \in E : f(x) \leq \alpha\}$$

is measurable. It follows from this that f is measurable.

□

3.2.3 Examples

[1] The constant function is measurable. That is, if (X, Σ) is a measurable space, $c \in \mathbb{R}$, and $E \in \Sigma$, then the function $f : E \to \mathbb{R}$ given by $f(x) = c$, for each $x \in E$, is measurable.

Let $\alpha \in \mathbb{R}$. If $\alpha \geq c$, then the set $\{x \in E : f(x) > \alpha\} = \emptyset$, and is thus measurable.

If $\alpha < c$, then the set $\{x \in E : f(x) > \alpha\} = E$, and is therefore measurable. Hence, f is measurable.

[2] Let (X, Σ) be a measurable space and let $A \in \Sigma$. The characteristic function, χ_A, is defined by

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

The characteristic function χ_A is measurable. This follows immediately from the observation that, for each $\alpha \in \mathbb{R}$ and $E \in \Sigma$, the set $\{x \in E : \chi_A > \alpha\}$ is either E, A or $\emptyset$.

[3] Let $\mathbb{B}$ be the Borel σ-algebra in $\mathbb{R}$ and $E \in \mathbb{B}$. Then, any continuous function $f : E \to \overline{\mathbb{R}}$ is measurable. This is an immediate consequence of the fact that, if $f : E \to \overline{\mathbb{R}}$ is continuous and $\alpha \in \mathbb{R}$, then the set $\{x \in E : f(x) > \alpha\}$ is open and hence belongs to $\mathbb{B}$.

3.2.4 Proposition

[1] If f and g are measurable real-valued functions defined on a common domain $E \in \Sigma$ and $c \in \mathbb{R}$, then the functions

(a) $f + c$,

(b) cf,

(c) $f \pm g$,

(d) f^2,

(e) $f \cdot g$,
(f) $|f|$,
(g) $f \vee g$,
(h) $f \wedge g$ are also measurable.

[2] If (f_n) is a sequence of measurable functions defined on a common domain $E \in \Sigma$, then the functions

(a) $\sup_n f_n$,
(b) $\inf_n f_n$,
(c) $\limsup_n f_n$,
(d) $\liminf_n f_n$ are also measurable.

PROOF.

[1] (a) For any real number α,

$$\{x \in E : f(x) + c > \alpha\} - \{x \in E : f(x) > \alpha - c\}.$$

Since the set on the right-hand side is measurable, we have that $f + c$ is measurable.

(b) If $c = 0$, then cf is obviously measurable. Assume that $c < 0$. Then, for each real number α,

$$\{x \in E : cf(x) > \alpha\} = \left\{x \in E : f(x) < \frac{\alpha}{c}\right\}.$$

Since the set $\{x \in E : f(x) < \frac{\alpha}{c}\}$ is measurable, it follows that cf is also measurable.

(c) Let α be a real number. Since the rationals are dense in the reals, there is a rational number r such that

$$f(x) < r < \alpha - g(x),$$

whenever $f(x) + g(x) < \alpha$. Therefore,

$$\{x \in E : f(x) + g(x) < \alpha\} = \bigcup_{r \in \mathbb{Q}} (\{x \in E : f(x) < r\} \cap \{x \in E : g(x) < \alpha - r\}).$$

Since the sets $\{x \in E : f(x) < r\}$ and $\{x \in E : g(x) < \alpha - r\}$ are measurable, so is the set $\{x \in E : f(x) < r\} \cap \{x : g(x) < \alpha - r\}$, and consequently, the set $\{x \in E : f(x) + g(x) < \alpha\}$, being a countable union of measurable sets, is also measurable.

If g is measurable, it follows from (b) that $(-1)g$ is also measurable. Hence, so is $f + (-1)g = f - g$.

(d) Let $\alpha \in \mathbb{R}$ and $E \in \Sigma$. If $\alpha < 0$, then $\{x \in E : f^2(x) > \alpha\} = E$, which is measurable.
If $\alpha \geq 0$, then

$$\{x \in E : f^2(x) > \alpha\} = \{x \in E : f(x) > \sqrt{\alpha}\} \cup \{x \in E : f(x) < -\sqrt{\alpha}\}.$$

Since the two sets on the right hand side are measurable, it follows that the set $\{x \in E : f^2(x) > \alpha\}$ is also measurable. Hence f^2 is measurable.

(e) Since $f \cdot g = \frac{1}{4}[(f+g)^2 - (f-g)^2]$, it follows from (b), (c), and (d) that $f \cdot g$ is measurable.

(f) Let $\alpha \in \mathbb{R}$ and $E \in \Sigma$. If $\alpha < 0$, then $\{x \in E : |f(x)| > \alpha\} = E$, which is measurable.
If $\alpha \geq 0$, then

$$\{x \in E : |f(x)| > \alpha\} = \{x \in E : f(x) > \alpha\} \cup \{x \in E : f(x) < -\alpha\}.$$

Since the two sets on the right hand side are measurable, it follows that the set $\{x \in E : f(x)^2 > \alpha\}$ is also measurable. Thus, $|f|$ is measurable.

(g) It is sufficient to observe that $f \vee g = \frac{1}{2}\{f + g + |f - g|\}$. It now follows from (b), (c), and (d) that $f \vee g$ is measurable.

(h) It is sufficient to observe that $f \wedge g = \frac{1}{2}\{f + g - |f - g|\}$. It now follows from (b), (c), and (d) that $f \wedge g$ is measurable.

[2] (a) Let $\alpha \in \mathbb{R}$. Then

$$\{x \in E : \sup_n f_n > \alpha\} = \bigcup_{n=1}^{\infty}\{x \in E : f_n(x) > \alpha\}.$$

Since for each $n \in \mathbb{N}$, f_n is measurable, it follows that the set $\{x \in E : f_n(x) > \alpha\}$ is measurable for each $n \in \mathbb{N}$. Therefore, the set $\{x \in E : \sup_n f_n(x) > \alpha\}$ is measurable as it is a countable union of measurable sets.

(b) It is sufficient to note that $\inf_n f_n = -\sup_n(-f_n)$.

(c) Notice that $\limsup_n f_n = \inf_{n\geq 1}\left(\sup_{k\geq n} f_k\right)$ and use (a) and (b) above.

(d) Notice that $\liminf_n f_n = \sup_{n\geq 1}\left(\inf_{k\geq n} f_k\right)$ and use (a) and (b) above.

□

3.2.5 Corollary

Let (X, Σ) be a measurable space. If f is a pointwise limit of a sequence (f_n) of measurable functions defined on a common domain $E \in \Sigma$, then f is measurable.

PROOF.

If the sequence (f_n) converges pointwise to f, then, for each $x \in E$,

$$f(x) = \lim_{n\to\infty} f_n(x) = \limsup_n f_n(x).$$

Now, by Proposition 3.2.4 [2], f is measurable.

□

3.2.6 Definition

Let f be a real-valued function defined on a set X. The **positive part of** f, denoted by f^+, is the function $f^+ = \max\{f, 0\} = f \vee 0$ and the **negative part of** f, denoted by f^-, is the function $f^- = \max\{-f, 0\} = (-f) \vee 0$.

Immediately we have that if f is a real-valued function defined on X, then

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

Note that

$$f^+ = \frac{1}{2}[|f| + f] \quad \text{and} \quad f^- = \frac{1}{2}[|f| - f].$$

Let (X, Σ) be a measurable space and $f : X \to \mathbb{R}$. It is trivial to deduce from Proposition 3.2.4 that f is measurable if and only if f^+ and f^- are measurable.

3.2.7 Definition

Let (X, Σ) be a measurable space. A **simple function** on X is a function of the form $\phi = \sum_{j=1}^{n} c_j \chi_{E_j}$, where, for each $j = 1, 2, \ldots n$, c_j is an extended real number and $E_j \in \Sigma$.

Since χ_{E_j} is measurable for each $j = 1, 2, \ldots, n$, it follows from Proposition 3.2.4 that ϕ is also measurable.

The representation of ϕ in Definition 3.2.7 is not unique. If ϕ assumes distinct nonzero values $a_1, a_2, \ldots, a_n$, then

$$\phi = \sum_{j=1}^{n} a_j \chi_{A_j},$$

where $A_j = \{x \in X : \phi(x) = a_j, j = 1, 2, \ldots, n\}$. This is called the **cannonical representation of ϕ**. Note that the sets A_j are pairwise disjoint and $X = \bigcup_{i=1}^{n} A_j$.

What Corollary 3.2.5 is saying is that the pointwise limit of a sequence of simple functions is a measurable function. The converse also holds. That is, every measurable function is a pointwise limit of a sequence of simple functions which we state as a theorem.

3.2.8 Theorem
Let (X, Σ) be a measurable space and f a nonnegative measurable function. Then there is a monotonic increasing sequence (ϕ_n) of nonnegative simple functions which converge pointwise to f.

PROOF.

For each $n \in \mathbb{N}$ and for $k = 1, 2, \ldots, n2^n$, let

$$E_{nk} = \left\{x \in X : \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}\right\}, \text{ and}$$
$$F_n = \{x \in X : f(x) \geq n\}.$$

Then $\{E_{nk} : k = 1, 2, \ldots n2^n\} \cup \{F_n : n \in \mathbb{N}\}$ is a collection of pairwise disjoint measurable sets whose union is X.

For each $n \in \mathbb{N}$, define ϕ_n by

$$\phi_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \chi_{E_{nk}} + n\chi_{F_n}.$$

Then the ϕ_n are nonnegative simple functions, $\phi_n \leq f$, and $\phi_n \leq \phi_{n+1}$, for each $n \in \mathbb{N}$. If $x \in X$ is such that $f(x) < \infty$, then, if $f(x) < n$,

$$0 \leq f(x) - \phi_n(x) < \frac{k}{2^n} - \frac{(k-1)}{2^n} = \frac{1}{2^n}.$$

Since $\lim_{n\to\infty} \frac{1}{2^n} = 0$, we have that $\lim_{n\to\infty} \phi_n(x) = n$, and so $\lim_{n\to\infty} \phi_n(x) = f(x)$.

□

3.2.9 Corollary
Let (X, Σ) be a measurable space and f a measurable function. Then there is a sequence (ϕ_n) of simple functions which converge pointwise to f.

PROOF.

The functions f^+ and f^- are both nonnegative and measurable. By Theorem 3.2.8, there exist two sequences (ϕ_n) and (ψ_n) of nonnegative simple functions which converge pointwise to f^+ and f^- respectively. The sequence (ζ_n), where, for each $n \in \mathbb{N}$, $\zeta_n = \phi_n - \psi_n$, is a sequence of simple functions which converges pointwise to $f^+ - f^- = f$.

□

3.2.10 Definition
Let (X, Σ, μ) be a measure space and $E \in \Sigma$. A property $P(x)$, where $x \in E$, is said to hold **almost everywhere** (abbreviated, a.e.) on E if the set $N = \{x \in E : P(x) \text{ fails}\}$ belongs to Σ and $\mu(N) = 0$.

3.2.11 Examples
[1] $f = g$ a.e. if $\mu(\{x \in X : f(x) \neq g(x)\}) = 0$.

[2] $f \leq g$ a.e. if $\mu(\{x \in X : f(x) > g(x)\}) = 0$.

3.2.12 Definition
Let (X, Σ, μ) be a measure space. A function $f : X \to \overline{\mathbb{R}}$ is said to be **almost everywhere real-valued**, denoted by a.e. real-valued, if

$$\mu(\{x \in X : |f(x)| = \infty\}) = 0.$$

We call a set of measure zero a **null set**.

3.2.13 Definition
A measure space (X, Σ, μ) is said to be **complete** if Σ contains all subsets of sets of measure zero. That is, if $E \in \Sigma$, $\mu(E) = 0$ and $A \subset E$, then $A \in \Sigma$.

It follows from Theorem 2.3.6 [1] that if a measure space (X, Σ, μ) is complete, $E \in \Sigma$, $\mu(E) = 0$ and $A \subset E$, then $\mu(A) = 0$.

3.2.14 Proposition
Let (X, Σ, μ) be a complete measure space and $f = g$ a.e. If f is measurable on $E \in \Sigma$, then so is g.

PROOF.

Let $\alpha \in \mathbb{R}$ and $N = \{x \in E : g(x) \neq f(x)\}$. Then $N \in \Sigma$ and $\mu(N) = 0$. Now,

$$\begin{aligned}\{x \in E : g(x) > \alpha\} &= \{x \in E \backslash N : g(x) > \alpha\} \cup \{x \in N : g(x) > \alpha\} \\ &= \{x \in E \backslash N : f(x) > \alpha\} \cup \{x \in N : g(x) > \alpha\}.\end{aligned}$$

The first set on the right hand side is measurable since f is measurable. The second set is measurable since it is a subset of the null set N and the measure is complete.

$\square$

3.2.2 Convergence of sequences of measurable functions

We now consider various notions of convergence of a sequence of measurable functions examine the relationships that exist between them.

3.2.15 Definition
Let $(X \Sigma, \mu)$ be a measure space. A sequence (f_n) of a.e. real-valued measurable functions on X is said to

[1] **converge almost everywhere** to an a.e. real-valued measurable function f. denoted by $f_n \to^{\text{a.e.}} f$, if for each $\epsilon > 0$ and each $x \in X$, there is a set $E \in \Sigma$ and a natural number $N = N(\epsilon)$ such that $\mu(E) < \epsilon$ and

$$|f_n(x) - f(x)| < \epsilon, \text{ for each } x \in X \backslash E \text{ and each } n \geq N.$$

[2] **converge almost uniformly** to an a.e. real-valued measurable function f, denoted by $f_n \to^{\text{a.u.}} f$, if for each $\epsilon > 0$, there is a set $E \in \Sigma$ and a natural number $N = N(\epsilon)$ such that $\mu(E) < \epsilon$ and

$$\|f_n - f\|_\infty = \sup_{x \in X \backslash E} |f_n(x) - f(x)| < \epsilon, \text{ for each } n \geq N.$$

[3] **converge in measure** to an a.e. real-valued measurable function f, denoted by $f_n \to^{\mu} f$, if for every $\epsilon > 0$,

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

(if μ is a probability measure, then this mode of convergence is called **convergence in probability**).

3.2.16 Proposition
Let (X, Σ, μ) be a complete measure space and (f_n) a sequence of measurable functions on $E \in \Sigma$ which converges to f a.e. Then f is measurable on E.

PROOF.

Let $\alpha \in \mathbb{R}$ and $N = \{x \in E : \lim_{n\to\infty} f_n(x) \neq f(x)\}$. Then $N \in \Sigma, \mu(N) = 0$ and $\lim_{n\to\infty} f_n(x) = f(x)$, for each $x \in E\backslash N$. Since E and N are measurable, so is $E\backslash N$. For each $n \in \mathbb{N}$, define g_n and g by

$$g_n(x) = \begin{cases} 0 & \text{if } x \in N \\ f_n(x) & \text{if } x \in E\backslash N. \end{cases}$$

and

$$g(x) = \begin{cases} 0 & \text{if } x \in N \\ f(x) & \text{if } x \in E\backslash N. \end{cases}$$

Then $g_n = f_n$ a.e. and $f = g$ a.e. It follows from Proposition 3.2.4, that for each $n \in \mathbb{N}$, g_n is measurable. If $x \in N$, then

$$\lim_{n\to\infty} g_n(x) = \lim_{n\to\infty} 0 = 0 = g(x), \tag{3.1}$$

and if $x \in E\backslash N$, then

$$\lim_{n\to\infty} g_n(x) = \lim_{n\to\infty} f_n = 0 = f(x) = g(x) \tag{3.2}$$

It follows from (3.1) and (3.2) that the sequence (g_n) converges pointwise to g *everywhere* on E. By Corollary 3.2.5, that g is measurable. Since $g = f$ a.e., we have, by Proposition 3.2.14, that f is measurable.

□

3.2.17 Theorem
Let (X, Σ, μ) be a measure space and (f_n) be a sequence of real-valued measurable functions on X. If the sequence (f_n) converges almost uniformly to f, then it converges in the measure to f. That is, almost uniform convergence implies convergence in the measure.

PROOF.

Let $\epsilon > 0$ be given. Since the sequence (f_n) converges almost uniformly to f, there is a measurable set E and a natural number N such that $\mu(E) < \epsilon$ and

$$|f_n(x) - f(x)| < \epsilon, \text{ for all } x \in X\backslash E \text{ and all } n \geq N.$$

It now follows that, for all $n \geq N, \{x \in X : |f_n(x) - f(x)| \geq \epsilon\} \subset E$. Therefore, for all $n \in \mathbb{N}$,

$$\mu(\{x \in X : |f_n(x) - f(x)| \geq \epsilon\}) < \epsilon.$$

□

3.2.18 Theorem
Let (X, Σ, μ) be a measure space and (f_n) be a sequence of a.e. real-valued measurable functions on X. If the sequence (f_n) converges almost uniformly to f, then it converges almost everywhere to f. That is, almost uniform convergence implies convergence almost everywhere.

PROOF.

Suppose that the sequence (f_n) converges almost uniformly to f. Then, for each $n \in \mathbb{N}$, there is a measurable set E_n, with $\mu(E_n) < \frac{1}{n}$ such that $f_n \to f$ uniformly on $X\backslash E_n$. Let $A = \bigcup_{n=1}^{\infty}(X\backslash E_n)$. Then

$$\mu(X\backslash A) = \mu\left(\bigcap_{n=1}^{\infty} E_n\right) \leq \mu(E_n) = \frac{1}{n} \to 0.$$

That is, $\mu(X\backslash A) = 0$. Furthermore, for each $x \in A$, $f_n(x) \to f(x)$ as $n \to \infty$. Hence, $f_n \to^{\text{a.e.}} f$.

□

In general, the converses of the previous two theorems do not hold. The following is an example of a sequence that converges almost everywhere and in measure, but does not converge almost uniformly.

3.2.19 Example

Let $X = [0, \infty)$ with the Lebesgue measure μ. For each $n \in \mathbb{N}$, let $f_n = \chi_{[n,n+\frac{1}{n}]}$, i.e.,

$$f_n(x) = \begin{cases} 1 & \text{if } x \in [n, n + \frac{1}{n}] \\ 0 & \text{otherwise.} \end{cases}$$

Then, for each $x \in X$, $f_n(x) \to^{n\to\infty} 0$. That is, the sequence (f_n) converges pointwise on X to the zero function. Since $\{x \in X : f_n(x) \not\to 0\} = \emptyset$, it follows that $\mu(\{x \in X : f_n(x) \not\to 0\}) = 0$, Hence, $f_n \to^{\text{a.e.}} 0$.

The sequence (f_n) converges in measure to the zero function. Indeed, for each $\epsilon > 0$,

$$\mu(\{x \in X : |f_n(x) - 0| \geq \epsilon\}) = \frac{1}{n} \to^{n\to\infty} 0.$$

We show that the sequence (f_n) does not converge almost uniformly to the zero function. Assume, on the contrary that (f_n) converges almost uniformly to the zero function. Then, there is a Lebesgue measurable set E with $\mu(E) = 0$ such that $f_n \to 0$ such that $f_n \to 0$ uniformly on E^c. Therefore, with $\epsilon = 1$, there is an $N \in \mathbb{N}$ such that for all $x \in E^c$ and all $n \geq N$,

$$\|f_n - 0\|_\infty = \|f_n\|_\infty < 1 \Leftrightarrow |f_n(x)| = f_n(x) < 1.$$

For each $n \in \mathbb{N}$, let $A_n = [n, n + \frac{1}{n}]$. Then, $E^c \cap \bigcup_{n\geq N} A_n = \emptyset$, and so $\bigcup_{n\geq N} A_n \subset E$. Thus,

$$0 = \mu(E) \geq \mu\left(\bigcup_{n\geq N} A_n\right) = \sum_{n=N}^{\infty} \mu(A_n) = \sum_{n=N}^{\infty} \frac{1}{n} = \infty,$$

which is absurd. Therefore (f_n) does not converge almost uniformly to the zero function.

3.2.20 Theorem (Egoroff's Theorem)

Let (X, Σ, μ) be a finite measure space and (f_n) be a sequence of a.e. real-valued measurable functions on X. If the sequence (f_n) converges almost everywhere to f, then it converges almost uniformly to f.

PROOF.

Without loss of generality, we assume that (f_n) converges to f everywhere on X. For $m, n \in N$, let

$$E_n^m = \bigcup_{k=n}^{\infty} \{x \in X : |f_k(x) - f(x)| \geq \frac{1}{m}\}.$$

Then, for $m, n \in \mathbb{N}$, E_n^m is a measurable set and $E_{n+1}^m \subset E_n^m$. That is, $(E_n^m)_n$ is a decreasing sequence of measurable subsets of X. Since $f_n(x) \to^{n\to\infty} f(x)$ for each $x \in X$, we have that $\bigcap_{n=1}^{\infty} E_n^m = \emptyset$. Also, since $\mu(X) < \infty$, we have that $\mu(E_1^m) < \infty$. By Theorem 2.3.6 [4], it follows that

$$0 = \mu(\emptyset) = \mu\left(\bigcap_{n=1}^{\infty} E_n^m\right) = \lim_{n\to\infty} \mu(E_n^m).$$

For $\delta > 0$, let $k_m \in \mathbb{N}$ be such that $\mu(E^m_{k_m}) < \frac{\delta}{2^m}$. Let $E_\delta = \bigcup_{m=1}^{\infty} E^m_{k_m}$. Then, E_δ is a measurable set and

$$\mu(E_\delta) \leq \sum_{m=1}^{\infty} \mu(E^m_{k_m}) < \delta.$$

Notice that

$$\begin{aligned} X\backslash E_\delta = X\backslash \bigcup_{m=1}^{\infty} E^m_{k_m} &= \bigcap_{m=1}^{\infty} (X\backslash E^m_{k_m}) \\ &= \bigcap_{m=1}^{\infty} \left[X\backslash \bigcup_{k=k_m}^{\infty} \left\{ x \in X : |f_k(x) - f(x)| \geq \frac{1}{m} \right\} \right] \\ &= \bigcap_{m=1}^{\infty} \bigcap_{k=k_m}^{\infty} \left\{ x \in X : |f_k(x) - f(x)| < \frac{1}{m} \right\}. \end{aligned}$$

Let $\epsilon > 0$. Choose $m \in \mathbb{N}$ such that $\frac{1}{m} < \epsilon$. Then for any $x \in X\backslash E_\delta$ and any $k \geq k_m$,

$$|f_k(x) - f(x)| < \frac{1}{m} < \epsilon.$$

That is, (f_n) converges uniformly to f on $X\backslash E_\delta$.

□

3.2.21 Corollary
Let (X, Σ, μ) be a finite measure space and (f_n) be a sequence of a.e. real-valued measurable functions on X. If the sequence (f_n) converges almost everywhere to f, then it converges in measure to f.

PROOF.

This follows from Theorem 3.2.17 and Theorem 3.2.20.

□

The hypothesis that $\mu(X) < \infty$ in Egoroff's Theorem is essential. Let $X = \mathbb{R}$ with the Lebesgue measure μ. For each $n \in \mathbb{N}$, let $f_n = \chi_{[n,\infty)}$, i.e.,

$$f_n(x) = \begin{cases} 1 & \text{if } x \in [n, \infty) \\ 0 & \text{otherwise.} \end{cases}$$

Then, for each $x \in \mathbb{R}$, $f_n(x) \to^{n\to\infty} 0$. That is, the sequence (f_n) converges pointwise on $\mathbb{R}$ to the zero function. Since $\{x \in \mathbb{R} : f_n(x) \not\to 0\} = \emptyset$, it follows that $\mu(\{x \in \mathbb{R} : f_n(x) \not\to 0\}) = 0$. Thus, $f_n \to^{\text{a.e.}} 0$. We show that the sequence (f_n) does not converge almost uniformly to the zero function. Assume, on the contrary, that (f_n) converges almost uniformly to the zero function. Then, there is a Lebesgue measurable set E with $\mu(E) = 0$ such that $f_n \to 0$ uniformly on E^c. Therefore, with $\epsilon = 1$, there is an $N \in \mathbb{N}$ such that for all $x \in E^c$ and all $n \geq N$,

$$\|f_n - 0\|_\infty = \|f_n\|_\infty < 1 \Leftrightarrow |f_n(x)| = f_n(x) < 1.$$

For each $n \in \mathbb{N}$, let $A_n = [n, \infty)$. Then, $E^c \cap \bigcup_{n\geq N} A_n = E^c \cap A_N = \emptyset$, and so $A_N \subset E$.

Hence, $0 = \mu(E) \geq \mu(A_N) = \infty$, which is a contradiction. Therefore (f_n) does not converge almost uniformly to the zero function.

3.2.22 Theorem (F. Riesz)
Let (X, Σ, μ) be a measurable space and (f_n) be a sequence of a.e. real-valued measurable functions on X. If the sequence (f_n) converges in measure to f, then there is a subsequence (f_{n_k}) of (f_n) which converges almost everywhere to f.

PROOF.

Let $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$,

$$\mu(\{x \in X : |f_n(x) - f(x)| \geq 1\}) < \frac{1}{2}.$$

Now, choose $n_2 \in \mathbb{N}$ such that $n_2 > n_1$ and for all $n \geq n_2$,

$$\mu\left(\{x \in X : |f_n(x) - f(x)| \geq \frac{1}{2}\}\right) < \frac{1}{2^2}.$$

Next, choose $n_3 \in \mathbb{N}$ such that $n_3 > n_2$ and for all $n \geq n_3$,

$$\mu\left(\{x \in X : |f_n(x) - f(x)| \geq \frac{1}{3}\}\right) < \frac{1}{2^3}.$$

Continue this process obtaining an increasing sequence (n_k) of natural numbers with

$$\mu\left(\{x \in X : |f_{n_k}(x) - f(x)| \geq \frac{1}{k}\}\right) < \frac{1}{2^k}$$

For each $k \in \mathbb{N}$, let

$$A_k = \{x \in X : |f_{n_k}(x) - f(x)| \geq \frac{1}{k}\}.$$

Since, for each $k \in \mathbb{N}$, $\mu(A_k) < \frac{1}{2^k}$, we have that the series $\sum_{k=1}^{\infty} \mu(A_k)$ converges. By the Borel-Cantelli Lemma (Theorem 2.3.7), $\mu(\limsup A_k) = 0$. Let

$$A = \limsup A_k = \bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} A_j.$$

Choose $x \in X \setminus A$. Then there is an index $j_x \in \mathbb{N}$ such that $x \in X \setminus A_{j_x}$, Note that

$$\begin{aligned} X \setminus A_{j_x} &= X \setminus \bigcup_{k=j_x}^{\infty} \{x \in X : |f_{n_k}(x) - f(x)| \geq \frac{1}{k}\} \\ &= \bigcap_{k=j_x}^{\infty} \left[X \setminus \{x \in X : |f_{n_k}(x) - f(x)| \geq \frac{1}{k}\} \right] \\ &= \bigcap_{k=j_x}^{\infty} \{x \in X : |f_{n_k}(x) - f(x)| \geq \frac{1}{k}\}. \end{aligned}$$

Hence, for all $k \geq j_x$, $|f_{n_k}(x) - f(x)| < \frac{1}{k}$. Let $\epsilon > 0$ be given. Choose $k_0 \geq j_x$ such that $\frac{1}{k_0} < \epsilon$. Then, for all $k \geq k_0$,

$$|f_{n_k}(x) - f(x)| < \frac{1}{k} < \epsilon.$$

Therefore, for each $x \in X \setminus A$, $f_{n_k}(x) \to f(x)$, and hence, $f_{n_k} \to^{\text{a.e.}} f$.

□

3.3 Definition of the integral

In this section, we define a Lebesgue integral. This is done in three stages: firstly, we define the integral of a nonnegative simple function; then, using the integral of a nonnegative simple function, we define the integral or any measurable function. We also prove three key results: Monotone Convergence Theorem, Dominated Convergence Theorem and Fatou's Lemma.

Unless otherwise specified, we will work in the measure space (X, Σ, μ).

3.3.1 Integral of a nonnegative simple function

3.3.1 **Definition**
Let ϕ be a nonnegative simple function with the canonical representation $\phi = \sum_{i=1}^{n} a_i \chi_{E_i}$. The **Lebesgue integral of ϕ with respect to** μ, denoted by $\int_X \phi d\mu$, is the extended real number

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

If $A \in \Sigma$, we define

$$\int_A \phi d\mu = \int_X \chi_A \phi d\mu.$$

The function ϕ is integrable if $\int_X \phi d\mu$ is finite.

We also write $\int \phi d\mu$ for $\int_X \phi d\mu$.

It is straightforward to show that if A is a measurable set and ϕ is as above, then

$$\int_A \phi d\mu = \sum_{i=1}^{n} a_i \mu(A \cap E_i).$$

Furthermore,

$$\int_A d\mu = \int_A \chi_A d\mu = \int_X \chi_A d\mu = \mu(A).$$

It is necessary to show that the above definition of the Lebesgue integral is unambiguous, i.e, the integral is independent of the representation of ϕ. Assume that

$$\phi = \sum_{i=1}^{n} a_i \chi_{E_i} = \sum_{j=1}^{m} b_j \chi_{F_i},$$

where $E_i \cap E_k = \emptyset$, for all $1 \leq i \neq k \leq n$, $X = \bigcup_{i=1}^{n} E_i$; $F_k \cap F_l = \emptyset$, for all $1 \leq j \neq l \leq m$, and $\bigcup_{j=1}^{m} F_j$. Then for each $i = 1, 2, \ldots, n$ and each $j = 1, 2, \ldots, m$,

$$E_i = E_i \cap X = E_i \cap \left(\bigcup_{j=1}^{m} F_j \right) = \bigcup_{j=1}^{m} (E_i \cap F_j), \text{ a disjoint union, and}$$

$$F_j = F_j \cap X = \left(\bigcup_{i=1}^{n} E_i \right) = \bigcup_{i=1}^{n} (E_i \cap F_j), \text{ a disjoint union.}$$

If $E_i \cap F_j \neq \emptyset$, then, for $x \in E_i \cap F_j$, $a_i = \phi(x) = b_j$. That is, $a_i = b_j$ in this case. If $E_i \cap F_j = \emptyset$, then $\mu(E_i \cap F_j) = 0$. Thus, $a_i\mu(E_i \cap F_j) = b_j\mu(E_i \cap F_j)$, for all $1 \le i \le n$ and for all $1 \le j \le m$. Therefore

$$\begin{aligned}\sum_{i=1}^{n} a_i\mu(E_i) &= \sum_{i=1}^{n} a_i\mu\Big(\bigcup_{j=1}^{m}(E_i \cap F_j)\Big) = \sum_{i=1}^{n} a_i \sum_{j=1}^{m}\mu(E_i \cap F_j)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\mu(E_i \cap F_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} b_j\mu(E_i \cap F_j)\\ &= \sum_{j=1}^{m} b_j \sum_{i=1}^{n}\mu(E_i \cap F_j) = \sum_{j=1}^{m} b_j\mu\Big(\bigcup_{i=1}^{n}(E_i \cap F_j)\Big)\end{aligned}$$

$$\sum_{j=1}^{m} b_j\mu(F_j).$$

Hence the definition of the integral of ϕ is unambiguous.

3.3.2 Proposition

Let (X, Σ, μ) be a measure space, ϕ and ψ nonnegative simple functions, and c a nonnegative real number. Then

[1] $\int_X (\phi + \psi)d\mu = \int_X \phi d\mu + \int_X \psi d\mu$.

[2] $\int_X c\phi d\mu = c\int_X \phi d\mu$.

[3] If $\phi \le \psi$, then $\int_X \phi d\mu \le \int_X \psi d\mu$.

[4] If A and B are disjoint measurable sets, then

$$\int_{A\cup B} \phi d\mu = \int_A \phi d\mu + \int_B \phi d\mu.$$

[5] The set function $\nu : \Sigma \to [0, \infty]$ defined by

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

for $A \in \Sigma$, is a measure on X.

[6] If $A \in \Sigma$ and $\mu(A) = 0$, then $\int_A \phi d\mu = 0$.

PROOF.

Let $\phi = \sum_{i=1}^{n} a_i\chi_{E_i}$, and $\psi = \sum_{j=1}^{m} b_j\chi_{F_j}$ be canonical representations of ϕ and ψ respectively. Then $E_i \cap E_k = \emptyset$, for all $i \le i \neq k \le n$, $F_j \cap F_l = \emptyset$, for all $1 \le j \neq l \le m$, $X = \cup_{i=1}^{n} E_i$, and $X = \cup_{j=1}^{m} F_j$.

[1] For $1 \le i \le n$ and $1 \le j \le m$, let $G_{ij} = E_i \cap F_j$. Then, for each $x \in G_{ij}$, $\phi(x) + \psi(x) = a_i + b_j$, i.e., the function $\phi + \psi$ takes the values $a_i + b_j$ on $E_i \cap F_j$. For all $1 \le i, k \le n$ and $1 \le j, l \le m$,

$$\begin{aligned}G_{ij} \cap G_{kl} &= (E_i \cap F_j) \cap (E_k \cap F_l)\\ &= (E_i \cap E_k) \cap (F_j \cap F_l)\\ &= \emptyset \cap \emptyset = \emptyset.\end{aligned}$$

That is, $\{G_{ij} : 1 \leq i \leq n \text{ and } 1 \leq j \leq m\}$ is a collection of nm pairwise disjoint sets. Furthermore,

$$E_i = E_i \cap X = E_i \cap \left(\bigcup_{j=1}^{m} F_j\right) = \bigcup_{j=1}^{m}(E_i \cap F_j) = \bigcup_{j=1}^{m} G_{ij},$$

$$F_j = F_j \cap X = F_j \cap \left(\bigcup_{i=1}^{n} E_i\right) = \bigcup_{i=1}^{n}(E_i \cap F_j) = \bigcup_{i=1}^{n} G_{ij}, \text{ and}$$

$$X = X \cap X = \left(\bigcup_{i=1}^{n} E_i\right) \cap \left(\bigcup_{j=1}^{m} F_j\right) = \bigcup_{i=1}^{n}\bigcup_{j=1}^{m}(E_i \cap F_j) = \bigcup_{i=1}^{n}\bigcup_{j=1}^{m} G_{ij}$$

Therefore,

$$\phi + \psi = \sum_{i=1}^{n}\sum_{j=1}^{m}(a_i + b_j)\chi_{G_{ij}}.$$

With this representation of $\phi + \psi$, the numbers $a_i + b_j$ are not necessarily distinct. Let $c_1, c_2, \ldots, c_k$ be the distinct values assumed by $\phi + \psi$. Then

$$\begin{aligned}
\int_X (\phi + \psi)d\mu &= \sum_{r=1}^{k} c_r \mu(\{x : \phi(x) + \psi(x) = c_r\}) \\
&= \sum_{r=1}^{k} c_r \mu\left(\bigcup_{a_i+b_j=c_r} E_i \cap F_j\right) \\
&= \sum_{r=1}^{k} c_r \sum_{a_i+b_j=c_r} \mu(E_i \cap F_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m}(a_i + b_j)\mu(G_{ij}) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\mu(G_{ij}) + \sum_{i=1}^{n}\sum_{j=1}^{m} b_j\mu(G_{ij}) \\
&= \sum_{i=1}^{n} a_i \sum_{j=1}^{m}\mu(G_{ij}) + \sum_{i=1}^{n} b_j \sum_{j=1}^{m}\mu(G_{ij}) \\
&= \sum_{i=1}^{n} a_i\mu(E_i) + \sum_{j=1}^{m} b_j\mu(F_j) \\
&= \int_X \phi d\mu + \int_X \psi d\mu.
\end{aligned}$$

[2] If $c = 0$, then $c\phi$ vanishes identically and hence, $\int_X c\phi d\mu = c\int_X \phi d\mu$. Assume that $c > 0$. If $\phi = \sum_{i=1}^{n} a_i\chi_{E_i}$ is the canonical representation of ϕ, then $c\phi = \sum_{i=1}^{n} ca_i\chi_{E_i}$ is the canonical representation of $c\phi$. Therefore,

$$\int_X c\phi d\mu = \sum_{i=1}^{n} ca_i\mu(E_i) = c\sum_{i=1}^{n} a_i\mu(E_i) = c\int_X \phi d\mu.$$

[3] As shown above,

$$\int_X \phi = \sum_{i=1}^{n} a_i \mu(E_i) = \sum_{i=1}^{n} a_i \mu\left(\bigcup_{j=1}^{m} E_i \cap F_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i \mu(E_i \cap F_j), \text{ and}$$

$$\int_X \psi = \sum_{j=1}^{m} b_j \mu(F_j) = \sum_{j=1}^{m} b_j \mu\left(\bigcup_{i=1}^{n} F_j \cap E_i\right) = \sum_{j=1}^{m}\sum_{i=1}^{n} b_j \mu(E_i \cap F_j).$$

Now suppose that $\phi \le \psi$. Then, for each $x \in E_i \cap F_j$, $a_i = \phi(x) \le \psi(x) = b_j$. i.e., $a_i \le b_j$, for all i, j such that $E_i \cap F_j \neq \emptyset$. It follows that $\int_X \phi d\mu \le \int_X \psi d\mu$.

[4] Let ϕ be as above. Then

$$\begin{aligned}
\int_{A\cup B} \phi d\mu &= \sum_{i=1}^{n} a_i \mu((A \cup B) \cap E_i) \\
&= \sum_{i=1}^{n} a_i \mu[(A \cap E_i) \cup (B \cap E_i)] \\
&= \sum_{i=1}^{n} a_i [\mu(A \cap E_i) + \mu(B \cap E_i)] \\
&= \sum_{i=1}^{n} a_i \mu(A \cap E_i) + \sum_{i=1}^{n} \mu(B \cap E_i) \\
&= \int_A \phi d\mu + \int_B \phi d\mu.
\end{aligned}$$

[5] We have that, for $A \in \Sigma$,

$$\nu(A) = \int_A \phi d\mu = \sum_{i=1}^{n} a_i \mu(A \cap E_i).$$

If $A = \emptyset$, then $A \cap E_i = \emptyset$, for each $i = 1, 2, \ldots, n$. Hence, $\mu(A \cap E_i) = 0$, for $i = 1, 2, \ldots, n$ and so $\nu(\emptyset) = 0$.

Let (A_k) be a sequence of pairwise disjoint measurable sets and $A = \bigcup_{k=1}^{\infty} A_k$. Then, for each $i = 1, 2, \ldots, n$,

$$E_i \cap A = E_i \cap \left(\bigcup_{k=1}^{\infty} A_k\right) = \bigcup_{k=1}^{\infty} (E_i \cap A_k), \quad \text{a disjoint union.}$$

Thus,

$$\begin{aligned}\nu(A) &= \sum_{i=1}^{n}\sum_{i=1}^{n} a_i\mu(A\cap E_i)\\ &= \sum_{i=1}^{n} a_i\mu\left(\bigcup_{k=1}^{\infty} E_i\cap A_k\right)\\ &= \sum_{i=1}^{n} a_i \sum_{k=1}^{\infty}\mu(E_i\cap A_k)\\ &= \sum_{k=1}^{\infty}\sum_{i=1}^{n} a_i\mu(E_i\cap A_k)\\ &= \sum_{k=1}^{\infty}\nu(A_k).\end{aligned}$$

That is, $\nu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty}\nu(A_k)$.

[6] We have by [5], that

$$0 \le \int_A \phi d\mu = \sum_{i=1}^{n} a_i\mu(A\cap E_i) \le \sum_{i=1}^{n} a_i\mu(A) = 0.$$

Thus, $\int_A \phi d\mu = 0$.

□

The next result asserts that changing a simple function on a null set does not change the integral.

3.3.3 Corollary
Let (X,Σ,μ) be a measure space and ϕ a nonnegative simple function. If A and B are measurable sets such that $A\subseteq B$ and $\mu(B\backslash A) = 0$, then

$$\int_A \phi d\mu = \int_B \phi d\mu.$$

PROOF.

Noting that $B = A\cup(B\backslash A)$, a disjoint union, we have by Proposition 3.3.2 [4] and [6], that

$$\int_B \phi d\mu = \int_{A\cup(B\backslash A)} \phi d\mu = \int_A \phi d\mu + \int_{B\backslash A}\phi d\mu = \int_A \phi d\mu + 0 = \int_A \phi d\mu.$$

□

3.3.2 Integral of a nonnegative measurable function

We showed in Theorem 3.2.8 that every nonnegative measurable function is a pointwise limit of an increasing sequence of of nonnegative simple functions. Using this fact, we are able to define the integral of a nonnegative measurable function using the integrals of nonnegative simple functions.

3.3.4 Definition
Let f be a nonnegative measurable function. The **Lebesgue integral** of f with respect to μ, denoted by $\int_X f d\mu$, is defined as

$$\int_X f d\mu = \sup\left\{\int_X \phi d\mu : 0 \leq \phi \leq f, \phi \text{ is a simple function}\right\}.$$

If $E \in \Sigma$, then we define the **Lebesgue integral of f over E with respect to** μ as

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

Notice that if f is a nonnegative simple function, then Definitions 3.3.1 and 3.3.4 coincide.

3.3.5 Proposition
Let f and g be nonnegative measurable functions and c a nonnegative real number.

[1] $\int_X c f d\mu = c \int_X f d\mu$.

[2] If $f \leq g$, then $\int_X f d\mu \leq \int_X g d\mu$.

[3] If A and B are measurable sets such that $A \subseteq B$, then

$$\int_A f d\mu \leq \int_B f d\mu.$$

PROOF.

[1] If $c = 0$, then the equality holds trivially. Assume that $c > 0$. Then

$$\begin{aligned}\int_X c f d\mu &= \sup\left\{\int_X \phi d\mu : 0 \leq \phi \leq c f, \phi \text{ a simple function}\right\}\\ &= \sup\left\{c\int_X \frac{\phi}{c} d\mu : 0 \leq \frac{\phi}{c} \leq f, \phi \text{ a simple function}\right\}\\ &= c\sup\left\{\int_X \frac{\phi}{c} d\mu : 0 \leq \frac{\phi}{c} \leq f, \phi \text{ a simple function}\right\}\\ &= c\int_X f \delta\mu.\end{aligned}$$

[2] Since $f \leq g$, it follows that $\{\phi : 0 \leq \phi \leq f, \phi \text{ a simple function}\} \subset \{\phi : 0 \leq \phi \leq g, \phi \text{ a simple function}\}$. Thus,

$$\int_X f d\mu = \sup_{0 \leq \phi \leq f} \int_X \phi d\mu \leq \sup_{0 \leq \phi \leq g} \int_X \phi d\mu = \int_X g d\mu.$$

[3] If $A \subseteq B$, then $\chi_A \leq \chi_B$. Therefore, for any nonnegative measurable function f, we have that $f\chi_A \leq f\chi_B$. By [2], it follows that

$$\int_A f d\mu = \int_X f\chi_A d\mu \leq \int_X f\chi_B d\mu = \in_B f d\mu.$$

□

3.3.6 Proposition (Chebyshev's inequality)
Let (X, Σ, μ) be a measure space, f a nonnegative measurable function and c a positive real number. Then

$$\mu(\{x \in X : f(x) \geq c\}) \leq \frac{1}{c} \int_X f d\mu.$$

PROOF.

Let $A = \{x \in X : f(x) \geq c\}$. Then, since $c \leq f$ on A and $A \subset X$,

$$\chi_A \leq \chi_A f \leq \chi_X f = f.$$

By Proposition 3.3.5, we have that

$$\begin{aligned} \int_X \chi_A c d\mu &\leq \int_X f d\mu \\ \Rightarrow c \int_X \chi_A d\mu &\leq \int_X f d\mu \\ \Rightarrow c \int_A d\mu &\leq \int_X f d\mu \\ \Rightarrow c\mu(A) &\leq \int_X f d\mu \\ \Rightarrow \mu(A) &\leq \frac{1}{c} \int_X f d\mu \end{aligned}$$

i.e., $\mu(\{x \in X : f(x) \geq c\}) \leq \frac{1}{c} \int_X f d\mu$.

□

3.3.7 Proposition
Let (X, Σ, μ) be a measure space and f a nonnegative measurable function on X. Then, $\int_X f d\mu = 0$ if and only if $f = 0$ a.e. on X.

PROOF.

Assume that $\int_X f d\mu = 0$. For each $n \in \mathbb{N}$, let $E_n = \left\{x \in X : f(x) \geq \frac{1}{n}\right\}$. Then by Chebyshev's inequality (Proposition 3.3.6),

$$0 \leq \mu(E_n) \leq n \int_X f d\mu = 0.$$

Hence, $\mu(E_n) = 0$ for each $n \in \mathbb{N}$. Since $E_n \subset E_{n+1}$, for each $n \in \mathbb{N}$ and

$$E = \{x \in X : f(x) > 0\} = \bigcup_{n=1}^{\infty} E_n,$$

it follows, by Theorem 2.3.6, that

$$\mu(E) = \mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n \to \infty} \mu(E_n) = 0.$$

That is, $f = 0$ a.e. on X.

Conversely, assume that $f = 0$ a.e. on X. Let $\phi = \sum_{i=1}^{n} a_i \chi_{E_i}$ be a nonnegative simple function, written in canonical form, such that $\phi \leq f$. Then $\phi = 0$ a.e. on X. Since $E_i = \{x \in X : \phi(x) = a_i, i = 1, 2, \ldots, n\}$, we have that $\mu\left(\bigcup_{i=1}^{n} E_i\right) = 0$. But since

$$\mu\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \mu(E_i),$$

we have that $\mu(E_i) = 0$, for each $i = 1, 2, \ldots, n$. Therefore,

$$\int_X \phi d\mu = \sum_{i=1}^{n} a_i \mu(E_i) = 0.$$

It then follows that

$$\int_X f d\mu = \sup_{0 \leq \phi \leq f} \int_X \phi d\mu = 0.$$

□

3.3.8 Proposition
Let (X, Σ, μ) be a measure space and f a nonnegative measurable function. If $A \in \Sigma$ and $\mu(A) = 0$, then

$$\int_A f d\mu = 0.$$

PROOF.

Let ϕ be a nonnegative simple function such that $\phi \leq f$. Then, by Proposition 3.3.2 [6], we have that $\int_A \phi d\mu = 0$. Therefore

$$\int_A f d\mu = \sup_{0 \leq \phi \leq f} \int_A \phi d\mu = \sup_{0 \leq \phi \leq f} \{0\} = 0.$$

□

3.3.9 Theorem (Monotone convergence theorem)
Let (X, Σ, μ) be a measure space, (f_n) a sequence of measurable functions on X such that $0 \leq f_n \leq f_{n+1}$, for every $n \in \mathbb{N}$ and $\lim_{n\to\infty} f_n(x) = f(x)$, for each $x \in X$. Then f is measurable and

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

PROOF.

We have already shown in Corollary 3.2.5 that f is measurable.
Since $0 \leq f_n \leq f_{n+1} \leq f$, for each $n \in \mathbb{N}$, it follows, by Proposition 3.3.5, that

$$\int_X f_n d\mu \leq \int_X f_{n+1} d\mu \leq \int_X f d\mu,$$

for each $n \in \mathbb{N}$. Hence,

$$\lim_{n\to\infty} \int_X f_n d\mu \leq \int_X f d\mu \tag{3.3}$$

We now prove the reverse inequality. Let ϕ be a simple function such that $0 \leq \phi \leq f$. Choose and fix ϵ such that $0 < \epsilon < 1$. For each $n \in \mathbb{N}$, let

$$A_n = \{x \in X : f_n(x) \geq (1-\epsilon)\phi(x)\}.$$

Now, for each $n \in \mathbb{N}$, A_n is measurable, and since (f_n) is an increasing sequence, we have that $A_n \subseteq A_{n+1}$. Furthermore $X = \bigcup_{n=1}^{\infty} A_n$. Since for each $n \in \mathbb{N}$, $A_n \subset X$, it follows that $\bigcup_{n=1}^{\infty} A_n \subseteq X$. For the reverse containment, let $x \in X$. If $f(x) = 0$, then $\phi(x) = 0$, and therefore $x \in A_1$. If on the other hand, $f(x) > 0$, then $(1-\epsilon)\phi(x) < f(x)$. Since $f_n \uparrow f$ and $0 < \epsilon < 1$, there is a natural number N such that $(1-\epsilon)\phi(x) \leq f_n(x)$, for all $n \geq N$. Thus, $x \in A_n$, for all $n \geq N$, and so, $X \subseteq \bigcup_{n=1}^{\infty} A_n$. Now, for each $n \in \mathbb{N}$,

$$\int_X f_n d\mu \geq \int_{A_n} f_n d\mu \geq (1-\epsilon) \int_{A_n} \phi d\mu. \tag{3.4}$$

Define $\nu : \Sigma \to [0, \infty]$ by

$$\nu(E) = \int_E \phi d\mu, \quad E \in \Sigma.$$

We have already shown in Proposition 3.3.2 [5], that ν is a measure on Σ. Since (A_n) is an increasing sequence of measurable sets such that $X = \bigcup_{n=1}^{\infty} A_n$, we have, by Theorem 2.3.6 [3], that

$$\nu(X) = \nu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \nu(A_n).$$

That is, $\int_X \phi d\mu = \lim_{n\to\infty} \int_{A_n} \phi d\mu$. Letting $n \to \infty$ in (3.4), we have that

$$\lim_{n\to\infty} \int_X f_n d\mu \geq (1-\epsilon) \int_X \phi d\mu.$$

Since ϵ is arbitrary, it follows that

$$\lim_{n\to\infty} \int_X f_n d\mu \geq \int_X \phi d\mu,$$

for any nonnegative simple function ϕ such that $\phi \leq f$. Then,

$$\lim_{n\to\infty} \int_X f_n d\mu \geq \sup\left\{\int_X \phi d\mu : \phi \text{ is simple and } 0 \leq \phi \leq f\right\} = \int_X f d\mu. \tag{3.5}$$

From (3.3) and (3.5), we have that

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

□

3.3.10 Corollary
Let f and g be nonnegative measurable functions. Then

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

PROOF.

By Theorem 3.2.8, there are monotonic increasing sequences (ϕ_n) and (ψ_n) of nonnegative simple functions such that $\phi_n \to f$ and $\psi_n \to g$, as $n \to \infty$. Then $(\phi_n + \psi_n)$ is a monotonic increasing sequence which converges to $f + g$. By the Monotone Convergence Theorem (Theorem 3.3.9), we have that

$$\begin{aligned}
\int_X (f+g)d\mu &= \lim_{n\to\infty} \int_X (\phi_n + \psi_n)d\mu \\
&= \lim_{n\to\infty} \left[\int_X \phi_n d\mu + \int_X \psi_n d\mu \right] \\
&= \lim_{n\to\infty} \int_X \phi_n d\mu + \lim_{n\to\infty} \int_X \psi_n d\mu \\
&= \int_X f d\mu + \int_X g d\mu \quad \text{(by the Monotone Convergence Theorem)}
\end{aligned}$$

□

3.3.11 Corollary
Let (X, Σ, μ) be a measure space and (f_n) a sequence of nonnegative measurable functions on X. Then, for each $n \in \mathbb{N}$,

$$\int_X \sum_{k=1}^{n} f_k d\mu = \sum_{k=1}^{n} \int_X f_k d\mu.$$

PROOF.

(Induction on n). If $n = 1$, then the result is obviously true. If $n = 2$, then, by Corollary 3.3.10, the result holds. Assume that

$$\int_X \sum_{k=1}^{n-1} f_k d\mu = \sum_{k=1}^{n-1} \int_X f_k d\mu.$$

then

$$\begin{aligned}
\int_X \sum_{k=1}^{n} f_k d\mu &= \int_X \left(\sum_{k=1}^{n-1} f_k + f_n \right) d\mu \\
&= \int_X \sum_{k=1}^{n-1} f_k + \int_X f_n d\mu \quad \text{(by Corollary 3.3.10)} \\
&= \sum_{k=1}^{n-1} \int_X f_k d\mu + \int_X f_n d\mu \quad \text{(by the induction hypothesis)} \\
&= \sum_{k=1}^{n} \int_X f_k d\mu.
\end{aligned}$$

□

3.3.12 Corollary (Beppo Levi)
Let (X, Σ, μ) be a measure space (f_n) a sequence of nonnegative measurable functions on X and $f =$

$\sum_{n=1}^{\infty} f_n$. Then f is measurable and

$$\int\limits_X f d\mu = \int\limits_X \Big(\sum_{n=1}^{\infty} f_n\Big) d\mu = \sum_{n=1}^{\infty} \int\limits_X f_n d\mu.$$

PROOF.

For each $m \in \mathbb{N}$, let $g_m = \sum_{n=1}^{m} f_n$. Then, for each $m \in \mathbb{N}$, g_m is measurable, $0 \leq g_m \leq g_{m+1}$ and $g_m \to f$ as $m \to \infty$. By the Monotone Convergence Theorem (Theorem 3.3.9), we have that

$$\begin{aligned}
\int\limits_X f d\mu &= \lim_{m\to\infty} \int\limits_X g_m d\mu \\
&= \lim_{m\to\infty} \int\limits_X \sum_{n=1}^{m} f_n d\mu \\
&= \lim_{m\to\infty} \sum_{n=1}^{m} \int\limits_X f_n d\mu \quad \text{(by Corollary 3.3.11)} \\
&= \sum_{n=1}^{\infty} \int\limits_X f_n d\mu
\end{aligned}$$

□

3.3.13 Corollary
Let (X, Σ, μ) be a measure space and f a nonnegative measurable function on X. If N is a null set in Σ, then

$$\int\limits_X f d\mu = \int\limits_{N^c} f d\mu.$$

PROOF.

By definition,

$$f\chi_N(x) = \begin{cases} f(x) & \text{if } x \in N \\ 0 & \text{if } x \in N^c. \end{cases}$$

That is, $f_{\chi_N} = 0$ a.e. By Proposition 3.3.7 and Corollary 3.3.10, we have that

$$\int\limits_X f d\mu = \int\limits_X (f\chi_N + f\chi_{N^c}) = \int\limits_X f\chi_N d\mu + \int\limits_X f\chi_{N^c} d\mu.$$

□

In the hypothesis if the Monotone Convergence Theorem (Theorem 3.3.9), we required that the monotone sequence of nonnegative measurable functions converge pointwise to the function f. We can weaken the 'pointwise convergence' to a.e. convergence and require that the limit be a measurable function.

3.3.14 Theorem (Monotone Convergence Theorem)
Let (X, Σ, μ) be a measure space, (f_n) a sequence of measurable functions on X such that $0 \leq f_n \leq f_{n+1}$, for every $n \in \mathbb{N}$ and $f_n \to^{\text{a.e.}} f$, where f is measurable. Then

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

PROOF.

Since $f_n \to^{\text{a.e.}}$, there is a set $N \in \Sigma$ such that $\mu(N) = 0$ and $f_n \to f$ pointwise on N^c. By Theorem 3.3.9,

$$\lim_{n\to\infty} \int_{N^c} f_n d\mu = \int_{N^c} f d\mu.$$

From Corollary 3.3.13, it follows that

$$\int_X f d\mu = \int_{N^c} f d\mu = \lim_{n\to\infty} \int_{N^c} f_n d\mu = \lim_{n\to\infty} \int_X f_n d\mu.$$

□

3.3.15 Proposition
Let (X, Σ, μ) be a measure space and f a nonnegative measurable function on X. Then the set function $\lambda : \Sigma \to [0, \infty]$ defined by

$$\lambda(A) = \int_A f d\mu, \quad A \in \Sigma$$

is a measure on X. Furthermore, for any nonnegative measurable function g,

$$\int_X g d\lambda = \int_X g f d\mu.$$

PROOF.

We note that

$$\lambda(A) = \int_A f d\mu = \int_X \chi_A f d\mu.$$

Since $f \geq 0$, it follows that $\lambda(A) \geq 0$, for each $A \in \Sigma$. If $A = \emptyset$, then $\chi_A = 0$ and so $\chi_A f = 0$. Therefore

$$\lambda(\emptyset) = \lambda(A) = \int_X \chi_A f d\mu = \int_X 0 d\mu = 0.$$

Let (A_n) be a sequence of pairwise disjoint measurable sets and $A = \bigcup_{n=1}^\infty A_n$. Then

$$\chi_A = \chi_{\bigcup_{n=1}^\infty A_n} = \sum_{n=1}^\infty \chi_{A_n}.$$

Therefore, $\chi_A f = \sum_{n=1}^\infty \chi_{A_n} f$, and so

$$\begin{aligned}
\lambda(A) = \int_X \chi_A f d\mu &= \int_X \sum_{n=1}^\infty \chi_{A_n} f d\mu \\
&= \sum_{n=1}^\infty \int_X \chi_{A_n} f d\mu \quad \text{(by Corollary 3.3.12)} \\
&= \sum_{n=1}^\infty \int_{A_n} f d\mu \\
&= \sum_{n=1}^\infty \lambda(A_n).
\end{aligned}$$

That is, $\lambda\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \lambda(A_n)$.

If $g = \chi_A$, for some $A \in \Sigma$, i.e., g is a characteristic function, then

$$\int_X g d\lambda = \int_X \chi_A d\lambda = \int_A d\lambda = \lambda(A)$$

and

$$\int_X g f d\mu = \int_X \chi_A f d\mu = \int_A f d\mu = \lambda(A).$$

Thus, $\int_X g d\lambda = \int_X g f d\mu$. Assume that $g = \phi$, a nonnegative simple function and let $\phi = \sum_{i=1}^{n} a_i \chi_{E_i}$ be the canonical representation of ϕ. Then,

$$\begin{aligned}
\int_X g d\lambda = \int_X \phi d\lambda = \int_X \left(\sum_{i=1}^{n} a_i \chi_{E_i}\right) d\lambda \\
&= \sum_{i=1}^{n} a_i \int_X \chi_{E_i} d\lambda \quad \text{(by Corollary 3.3.11 and Proposition 3.3.5)} \\
&= \sum_{i=1}^{n} a_i \int_{E_i} d\lambda \\
&= \sum_{i=1}^{n} a_i \lambda(E_i),
\end{aligned}$$

and

$$\begin{aligned}
\int_X g f d\mu = \int_X \phi f d\mu = \int_X \left(\sum_{i=1}^{n} a_i \chi_{E_i}\right) f d\mu \\
&= \sum_{i=1}^{n} a_i \int_X \chi_{E_i} f d\mu \quad \text{(by Corollary 3.3.11 and Proposition 3.3.5)} \\
&= \sum_{i=1}^{n} a_i \int_{E_i} f d\mu \\
&= \sum_{i=1}^{n} a_i \lambda(E_i).
\end{aligned}$$

Therefore, if g is a nonnegative simple function, then $\int_X g d\lambda = \int_X g f d\mu$.

Let g be a nonnegative measurable function. Then, by Theorem 3.2.8, there is an increasing sequence (ϕ_n) of nonnegative simple functions such that $\phi_n \to^{n\to\infty} g$. Then $\phi_n f \to^{n\to\infty} g f$. Clearly the $\phi_n f$ are nonnegative and measurable functions for each $n \in \mathbb{N}$ (see Proposition 3.2.4). By the Monotone Convergence Theorem (Theorem 3.3.9), we have that

$$\begin{aligned}
\int_X g d\lambda &= \lim_{n\to\infty} \int_X \phi_n d\lambda \\
&= \lim_{n\to\infty} \int_X \phi_n f d\mu \\
&= \int_X g f d\mu.
\end{aligned}$$

□

3.3.16 Theorem (Fatou's Lemma)
Let (X, Σ, μ) be a measure space and (f_n) a sequence of nonnegative measurable functions on X. Then

$$\int_X \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu.$$

PROOF.

For each $k \in \mathbb{N}$, let $g_k = \inf_{n\geq k} f_n$. Then each g_k is a nonnegative measurable function. Furthermore, for each $k \in \mathbb{N}$, $g_k \leq g_{k+1}$ and $\lim_{k\to\infty} g_k = \liminf_{n\to\infty} f_n$. By the Monotone Convergence Theorem (Theorem 3.3.9), we have that

$$\int_X \liminf_{n\to\infty} f_n d\mu = \lim_{k\to\infty} \int_X g_k d\mu. \tag{3.6}$$

Since $g_k \leq f_n$ for all $n \geq k$, it follows, by Proposition 3.3.5, that

$$\int_X g_k d\mu \leq \int_X f_n d\mu, \quad \text{(for all } n \geq k).$$

That is, the number $\int_X g_k d\mu$ is a lower bound for the sequence of numbers $\left(\int_X f_n d\mu\right)_{n\geq k}$. Hence,

$$\int_X g_k d\mu \leq \inf_{n\geq k} \int_X f_n d\mu,$$

and consequently

$$\lim_{k\to\infty} \int_X g_k d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu. \tag{3.7}$$

From (3.6) and (3.7), we have that

$$\int_X \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu.$$

□

3.3.3 Integrable functions

Let f be a measurable function on X. Recall that the positive part of f, denoted by f^+, is the nonnegative measurable function

$$f^+ = \max\{f, 0\},$$

and the negative part of f, denoted f^-, is the nonnegative measurable function

$$f^- = \max\{-f, 0\}.$$

We have already seen that

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

Suppose now that $f = g - h$, where g and h are nonnegative measurable functions. Then $f^+ \leq g$ and $f^- \leq h$. Indeed, if $f \geq 0$, then $f^+ = f$. Hence, for any $x \in X$,

$$f^+(x) = f(x) = g(x) - h(x) \leq g(x).$$

If $f < 0$, then $f^+ = 0$ and so, for any $x \in X$,

$$f^+(x) = 0 \leq g(x).$$

A similar argument shows that $f^- \leq h$.

3.3.17 Definition
Let (X, Σ, μ) be a measure space and f a measurable function on X. If $\int_X f^+ d\mu < \infty$ or $\int_X f^- d\mu < \infty$, then we define the **integral** of f, denoted by $\int_X f d\mu$, as the extended real number

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

If $\int_X f^+ d\mu < \infty$ and $\int_X f^- d\mu < \infty$, then f is said to be **(Lebesgue) integrable on** X. The set of all integrable functions is denoted by $\mathcal{L}^1(X, \mu)$.

It is clear that f is integrable if and only if $\int_X |f| d\mu < \infty$. In this case,

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

3.3.18 Remark
Let (X, Σ, μ) be a measure space and f a measurable function on X. If f is integrable on X, then $|f| < \infty$ a.e. on X. Indeed, if $A \in \Sigma$ with $\mu(A) > 0$ and $|f| = \infty$ on A, then, for each $n \in \mathbb{N}$, $|f| > n\chi_A$. Hence,

$$\int_X |f| d\mu \geq \int_X n\chi_A d\mu = n \int_A d\mu = n\mu(A).$$

Letting n tend to infinity, we have that $\int_X |f| d\mu = \infty$.

3.3.19 Theorem
Let (X, Σ, μ) be a measure space, $f, g \in \mathcal{L}^1(X, \mu)$ and let $c \in \mathbb{R}$. Then

[1] $cf \in \mathcal{L}^1(X, \mu)$ and $\int_X (cf) d\mu = c \int_X f d\mu$.

[2] $f + g \in \mathcal{L}^1(X, \mu)$ and $\int_X (f + g) d\mu = \int_X f d\mu + \int_X g d\mu$.

PROOF.

[1] Assume that $c \geq 0$. Then,

$$(cf)^+ = \max\{cf, 0\} = c \max\{f, 0\} = c(f^+) \text{ and}$$
$$(cf)^- = \max\{-(cf), 0\} = c \max\{-f, 0\} = c(f^-)$$

Therefore,

$$\int_X (cf)^+ d\mu = \int_X cf^+ d\mu = c \int_X f^+ d\mu < \infty \text{ and}$$
$$\int_X (cf)^- d\mu = \int_X cf^- d\mu = c \int_X f^- d\mu < \infty.$$

Hence, cf is integrable and

$$\begin{aligned}\int_X (cf)d\mu &= \int_X (cf)^+ d\mu - \int_X (cf)^- d\mu \\ &= c\int_X f^+ d\mu - c\int_X f^- d\mu \\ &= c\Big(\int_X f^+ d\mu - \int_X f^- d\mu\Big) \\ &= c\int_X f d\mu.\end{aligned}$$

If $c < 0$, then $c = -k$ for some $k > 0$. Therefore

$$(cf)^+ = \max\{cf, 0\} = \max\{-kf, 0\} = k \max\{-f, 0\} = kf^- = -cf^- \text{ and}$$
$$(cf)^- = \max\{-(cf), 0\} = \max\{-cf, 0\} = \max\{kf, 0\} = k \max\{f, 0\} = kf^+ = -cf^+.$$

From this it follows that

$$\int_X (cf)^+ d\mu = \int_X (-cf^-)d\mu = -c\int_X f^- d\mu < \infty \text{ and}$$
$$\int_X (cf)^- d\mu = \int_X (-cf^+)d\mu = -c\int_X f^+ d\mu < \infty.$$

Therefore, cf is integrable and

$$\begin{aligned}\int_X (cf)d\mu &= \int_X (cf)^+ d\mu - \int_X (cf)^- d\mu \\ &= -c\int_X f^- d\mu - (-c)\int_X f^+ d\mu \\ &= -c\int_X f^- d\mu + c\int_X f^+ d\mu \\ &= c\Big(\int_X f^+ d\mu - \int_X f^- d\mu\Big) \\ &= c\int_X f d\mu.\end{aligned}$$

[2] Since

$$f + g = (f + g)^+ - (f + g)^- \text{ and}$$
$$f + g = (f^+ - f^-) + (g^+ - g^-) = (f^+ g^+) - (f^- + g^-),$$

we have that

$$(f + g)^+ \leq f^+ g^+ \text{ and } (f + g)^- \leq f^- + g^-.$$

By Proposition 3.3.5 [2], it follows that

$$\int_X (f+g)^+ d\mu \le \int_X (f^+ + g^+) d\mu \le \int_X f^+ d\mu + \int_X g^+ d\mu < \infty \text{ and}$$
$$\int_X (f+g)^- d\mu \le \int_X (f^- + g^-) d\mu \le \int_X f^- d\mu + \int_X g^- d\mu < \infty.$$

Thus, $f+g$ is integrable. Since $(f+g)^+ - (f+g)^- = f+g = f^+ - f^- + g^+ - g^-$, we have that

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

By Corollary 3.3.10, it follows that

$$\begin{aligned}
&\int_X [(f+g)^+ + f^- + f^-] d\mu = \int_X [(f+g)^- + f^+ + g^+] d\mu \\
\Leftrightarrow &\int_X (f+g)^+ d\mu + \int_X f^- d\mu + \int_X g^- d\mu = \int_X (f+g)^- d\mu + \int_X f^+ d\mu + \int_X g^+ d\mu \\
\Leftrightarrow &\int_X (f+g)^+ d\mu - \int_X (f+g)^- d\mu = \int_X f^+ d\mu + \int_X f^- d\mu + \int_X g^+ d\mu - \int_X g^- d\mu \\
\Leftrightarrow &\int_X (f+g) d\mu = \int_X f d\mu + \int_X g d\mu.
\end{aligned}$$

□

3.3.20 Corollary
Let (X, Σ, μ) be a measure space. Then $\mathcal{L}^1(X, \mu)$ is a vector space with respect to the usual operations of addition and scalar multiplication.

3.3.21 Proposition
Let (X, Σ, μ) be a measure space, $f, g \in \mathcal{L}^1(X, \mu)$. If $f \le g$ a.e. on X, then

$$\int_X f d\mu \le \int_X g d\mu.$$

PROOF.

Since $g - f \ge 0$ a.e. on X, we have that

$$\int_X g d\mu - \int_X f d\mu = \int_X (g-f) d\mu \ge 0.$$

It now follows that $\int_X f d\mu \le \int_X g d\mu$.

□

3.3.22 Lemma
Let (X, Σ, μ) be a measure space. If g is a Lebesgue integrable function on X and f is a measurable function such that $|f| \le g$ a.e., then f is Lebesgue integrable on X.

PROOF.

Since $|f| = f^+ + f^-$, it follows that $f^+ \leq g$ and $f^- \leq g$. Hence,

$$0 \leq \int_X f^+ d\mu \leq \int_X g d\mu \text{ and } 0 \leq \int_X f^- d\mu \leq \int_X g d\mu.$$

Therefore f^+ and f^- are both integrable.

□

The following theorem, known as Lebesgue's Dominated Convergence Theorem, provides a useful criterion for the interchange of limits and integrals and is of fundamental importance in measure theory. As a motivation for the measure theoretic approach to integration, think about how careful one must be when interchanging the limits and integrals in Riemann integration.

3.3.23 Theorem (Lebesgue's Dominated Convergence Theorem)
Let (X, Σ, μ) be a measure space, (f_n) a sequence of measurable functions on X such that $f_n \to^{n\to\infty} f$ a.e., for f a measurable function. If there is a Lebesgue integrable function g such that for each $n \in \mathbb{N}$, $|f_n| \leq g$ a.e., then f is Lebesgue integrable and

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

PROOF.

Since $|f_n| \leq g$ a.e., for each $n \in \mathbb{N}$, it follows that $|f| \leq g$ a.e. By Lemma 3.3.22, we have that f is Lebesgue integrable.

Since $g \pm f_n \geq 0$, for each $n \in \mathbb{N}$, we have by Fatou's Lemma (Theorem 3.3.16), that

$$\int_X g d\mu + \int_X f d\mu = \int_X (g + f) d\mu \leq \liminf_n \int_X (g + f_n) d\mu, \text{ and} \tag{3.8}$$

$$\int_X g d\mu - \int_X f d\mu = \int_X (g - f) d\mu \leq \liminf_n \int_X (g - f_n) d\mu. \tag{3.9}$$

Now,

$$\liminf_n \int_X (g + f_n) d\mu = \int_X g d\mu + \liminf_n \int_X f_n d\mu, \text{ and} \tag{3.10}$$

$$\liminf_n \int_X (g - f_n) d\mu = \int_X g d\mu - \limsup_n \int_X f_n d\mu. \tag{3.11}$$

Since g is Lebesgue integrable, we have, from (3.8) and (3.10), that

$$\int_X f d\mu \leq \liminf_n \int_X f_n d\mu. \tag{3.12}$$

Similarly, from (3.9) and (3.11), we have that

$$\limsup_n \int_X f_n d\mu \leq \int_X f d\mu \tag{3.13}$$

From (3.12) and (3.13), it follows that

$$\limsup_n \int_X f_n d\mu \leq \int_X f d\mu \leq \liminf_n \int_X f_n d\mu,$$

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

□

Note that in the previous theorem, if we replace the almost everywhere convergence with pointwise convergence, then we can remove the assumption that f is measurable since f would be guaranteed to be measurable in that case. Furthermore, if we have a complete measure space (X, Σ, μ), then the almost everywhere limit of (f_n) is definitely measurable.

3.3.24 Remark

The existence of a Lebesgue-integrable dominating function g in Lebesgue's Dominated Convergence Theorem (Theorem 3.3.23) is essential. Let μ be the Lebesgue measure on $[0, 1]$ and, for each $n \in \mathbb{N}$, let $f_n = n\chi_{[0,\frac{1}{n}]}$. Then $f_n \to^{\text{a.e.}} 0$ and $f_n \to^{\mu} 0$, but $\int_0^1 f_n d\mu = 1 \neq 0$.

For another example, let μ again be the Lebesgue measure on $\mathbb{R}$. For each $n \in \mathbb{N}$, let $f_n = \frac{1}{n}\chi_{[0,n]}$. Then $f_n \to 0$ uniformly on $\mathbb{R}$, but $\int_{\mathbb{R}} f_n d\mu = 1 \neq 0$.

3.4 Lebesgue and Riemann integrals

In this section, we show that the Lebesgue integral is a generalization of the Riemann integral. the phrase 'f is Lebesgue-integrable on $[a, b]$' in this section means that f is integrable with respect to the Lebesgue measure on $[a, b]$.

3.4.1 Theorem

Let f be a bounded real-valued function on $[a, b]$. If f is Riemann-integrable on $[a, b]$ then f is Lebesgue-integrable on $[a, b]$ and the two integrals coincide.

PROOF.

Let $[a, b]$ be a closed and bounded interval of real numbers, f a bounded real-valued function on $[a, b]$ and $P = \{x_0, x_1, \ldots, x_n\}$ a partition of $[a, b]$. Let

$$M_i = \sup\{f(x) : x_{i-1} \leq x \leq x_i\}, \text{ and } m_i = \inf\{f(x) : x_{i-1} \leq x \leq x_i\},$$

for each $i = 1, 2, \ldots, n$. Recall that the upper sum of f relative to the partition P is

$$U(f, P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1})$$

and the lower sum of f relative to the partition P is

$$u(f, P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}).$$

Define $u(f, P) : [a, b] \to \mathbb{R}$ and $l(f, p) : [a, b] \to \mathbb{R}$ by

$$u(f, P) = M_i \text{ if } x_{i-1} \leq x \leq x_i, i = 1, 2, \ldots, n$$
$$l(f, P) = m_i \text{ if } x_{i-1} \leq x \leq x_i, i = 1, 2, \ldots, n.$$

Note that $u(f, P)$ and $l(f, P)$ are simple functions since they can be expressed in the form:

$$u(f, P) = \sum_{i=1}^{n} M_i\chi_{[x_{i-1},x_i]} \text{ and } l(f, P) = \sum_{i=1}^{n} m_i\chi_{[x_{i-1},x_i]}.$$

Let m be the Lebesgue measure on $[a, b]$. It is clear that the upper sum of f relative to the partition P, $U(f, P)$, is just the Lebesgue integral of the simple function $u(f, P)$ and the lower sum of f relative to the partition P, $L(f, P)$, is the Lebesgue integral of the function simple $l(f, P)$.

Let (P_n) be a sequence of partitions of $[a, b]$ such that, for each $n \in \mathbb{N}$, P_{n+1} is a refinement of P_n and $\max_i \triangle x_j = \max_j (x_i - x_{i-1})$ tends to zero as n tends to infinity. For each $n \in \mathbb{N}$, let $\phi_n = u(f, P_n)$ and $\psi_n = l(F, P_n)$. Then,

$$\psi_1 \le \psi_2 \le \cdots \le f \le \cdots \le \phi_2 \le \phi_1.$$

That is, the sequence (ϕ_n) is decreasing and bounded below, while the sequence (ψ_n) is increasing and bounded above. Hence, in the limit, $\psi \le \phi$, where $\phi = \lim_{n\to\infty} \phi_n = \inf_n \phi_n$ and $\psi \ \lim_{n\to\infty} \psi_n = \sup_n \psi_n$. Note that ϕ and ψ are measurable functions. If $|f|$ is bounded by M, then so are the functions $|\phi_n|$ and $|\psi_n|$, for each $n \in \mathbb{N}$. Since a constant function M is Lebesgue-integrable on $[a, b]$, it follows, by Lebesgue's Dominated Convergence Theorem (Theorem 3.3.23), that

$$\int_{[a,b]} \phi d\mu = \lim_{n\to\infty} \int_{[a,b]} \phi_n d\mu = \lim_{n\to\infty} U(f, P_n), \text{ and}$$

$$\int_{[a,b]} \psi d\mu = \lim_{n\to\infty} \int_{[a,b]} \psi_n d\mu = \lim_{n\to\infty} L(f, P_n)$$

Since f is Riemann-integrable on $[a, b]$, $\lim_{n\to\infty} U(f, P_n) = \int_a^b f(x)dx = \lim_{n\to\infty} L(f, P_n)$. This implies that

$$\int_{[a,b]} \psi d\mu = \int_{[a,b]} \phi d\mu = \int_{[a,b]} \phi d\mu \Leftrightarrow \int_{[a,b]} (\phi - \psi) d\mu = 0.$$

By Proposition 3.3.7, we have that $\phi = \psi$ a.e. with respect to the Lebesgue measure μ on $[a, b]$. Since $\psi \le f \le \phi$, it follows that $\psi = f = \phi$ a.e. $[\mu]$. It follows that f is Lebesgue integrable and

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

□

3.5 L^p spaces

Let (X, Σ, μ) be a measure space and $1 \le p < \infty$. Denote by $\mathcal{L}^p(X, \mu)$ the set of all measurable functions $f : X \to \overline{\mathbb{R}}$ such that $|f|^p$ is integrable; i.e.,

$$\mathcal{L}^p(X, \mu) = \{f : X \to \overline{R} : f \text{ is measurable and } \int_X |f|^p d\mu < \infty\}.$$

If $\alpha \in \mathbb{R}$ and $f \in \mathcal{L}^p(X, \mu)$, then αf is measurable and

$$\int_X |\alpha f|^p d\mu = |\alpha|^p \int_X |f|^p d\mu < \infty,$$

i.e., $\alpha f \in \mathcal{L}^p(X, \mu)$.

If $f, g \in \mathcal{L}^p(X, \mu)$, then $f + g$ is measurable and, since

$$|f + g|^p \le 2^{p-1}(|f|^p |g|^p),$$

it follows that

$$\int_X |f+g|^p d\mu \le \int_X 2^{p-1}(|f|^p + |g|^p)d\mu = 2^{p-1}\left(\int_X |f|^p d\mu + \int_X |g|^p d\mu\right) < \infty,$$

and so $f + g \in \mathcal{L}^p(X, \mu)$.

3.5.1 Definition
Let (X, Σ, μ) be a measure space. A measurable function $f : X \to \overline{\mathbb{R}}$ is said to be **essentially bounded** (with respect to μ) if there is a constant $M > 0$ such that

$$|f(x)| \le M \text{ a.e. on } X \tag{3.14}$$

The smallest constant M such that (3.14) holds is called the **essential supremum** of $|f|$ and is denoted by $\|f\|_\infty$; i.e.,

$$\|f\|_\infty = \text{ess sup}_X |f| := \inf\{M \in \mathbb{R} : |f| \le M \text{ a.e. on } X\}.$$

3.5.2 Remark
Note that f is essentially bounded if there is a constant $M > 0$ such that

$$\mu(\{x \in X : |f(x)| > M\}) = 0;$$

i.e., the sets of points in X where f fails to be bounded is a null set.

Denote by $\mathcal{L}^\infty(X, \mu)$ the set of all essentially bounded functions on a measure space $(X.\Sigma, \mu)$.

3.5.3 Proposition
Let (X, Σ, μ) be a measure space and $1 \le p \le \infty$. Then $\mathcal{L}^p(X, \mu)$ is a linear space.
PROOF.

Exercise.

□

3.5.1 Hölder and Minkowski inequalities

3.5.4 Definition
Let p and q be positive real numbers. If $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, or if $p = 1$ and $q = \infty$, or if $p = \infty$ and $q = 1$, then we say that p and q are **conjugate exponents**.

3.5.5 Lemma (Young's inequality)
Let p and q be conjugate exponents, with $1 < p, q < \infty$, $\alpha \ge 0$ and $\beta \ge 0$. Then

$$\alpha\beta \le \frac{\alpha^p}{p} + \frac{\beta^q}{q}.$$

$\prec$If $p = 2 = q$, then the inequality follows from the fact that $(\alpha - \beta)^2 \ge 0$.
Consider the function $f : [0, \infty) \to \mathbb{R}$ given by

$$f(\alpha) = \frac{\alpha^p}{p} + \frac{\beta^q}{q} - \alpha\beta, \text{ for fixed } \beta > 0.$$

Then $f(0) = \frac{\beta^q}{q} > 0$, $f(\beta^{q/p}) = 0$, and $\lim_{\alpha\to+\infty} f(\alpha) = +\infty$. Hence f has a global minimum on $[0, \infty)$, i.e., there is an $\alpha_0 \in [0, \infty)$ such that $f(\alpha_0) \le f(\alpha)$ for all $\alpha \in [0, \infty)$. Now, since $f'(\alpha_0) = 0$, it follows that

$$0 = f'(\alpha_0) = \alpha_0^{p-1} - \beta \iff \beta = \alpha_0^{p-1} \iff \alpha_0 = \beta^{\frac{1}{p-1}} = \beta^{q/p}.$$

Hence, for all $\alpha \in [0, \infty)$,

$$0 = f(\beta^{q/p}) \leq f(\alpha) = \frac{\alpha^p}{p} + \frac{\beta^q}{q} - \alpha\beta \iff \alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}. \qquad \blacksquare$$

3.5.6 Theorem (Hölder's inequality)
Let $1 \leq p, q \leq \infty$ be conjugate exponents, (X, Σ, μ) a measure space, $f \in \mathcal{L}^p(X, \mu)$ and $g \in \mathcal{L}^q(X, \mu)$. Then $fg \in \mathcal{L}^1(X, \mu)$ and

$$\int_X |fg| d\mu \leq \|f\|_p \|g\|_q.$$

PROOF.

Firstly, we consider the case that $1 < p < \infty$. If $\|f\|_p = 0$ or $\|g\|_q = 0$, then the result follows trivially. We therefore assume that $\|f\|_p > 0$ and $\|g\|_q > 0$. Let $F = \frac{f}{\|f\|_p}$ and $G = \frac{g}{\|g\|_q}$. Then, $F \in \mathcal{L}^p(X, \mu)$ and $G \in \mathcal{L}^q(X, \mu)$ and $\|F\|_p = 1\|G\|_q$. By Young's inequality (Lemma 3.5.5), we have that

$$|FG| \leq \frac{1}{p}|F|^p + \frac{1}{q}|G|^q.$$

It follows that

$$\int_x |FG| d\mu \leq \frac{1}{p} \int_X |F|^p d\mu + \frac{1}{q} \int_X |G|^q d\mu = \frac{1}{p} + \frac{1}{q} = 1,$$

and so,

$$\int_X |fg| d\mu \leq \|f\|_p \|g\|_q.$$

Now assume that $p = 1$ and $q = \infty$. Let $M = \|g\|_\infty$. Then $|fg| \leq M|f|$ a.e. on X and so

$$\int_X |fg| d\mu \leq \int_X M|f| d\mu \leq M \int_X |f| d\mu = \|g\|_\infty \|f\|_1.$$

$\square$

3.5.7 Theorem
Let (X, Σ, μ) be a finite measure space and $1 \leq r < p \leq \infty$. Then $\mathcal{L}^p(X, \mu) \subseteq \mathcal{L}^r(X, \mu)$.

PROOF.

Let $t = \frac{p}{r}$ and $s = \frac{p}{p-r}$. Then $\frac{1}{t} + \frac{1}{s} = 1$; i.e., s and t are conjugate exponents. Let $f \in \mathcal{L}^p(X, \mu)$. Then, by Hölder's inequality (Theorem 3.5.6),

$$\begin{aligned}
\|f\|_r^r = \int_X |f|^r d\mu &\leq \left(\int_X (|f|^r)^t d\mu \right)^{\frac{1}{t}} \left(\int_X 1^s d\mu \right)^{\frac{1}{s}} \\
&= \left(\int_X (|f|^r)^{\frac{p}{r}} d\mu \right)^{\frac{r}{p}} \left(\int_X 1^s d\mu \right)^{\frac{1}{s}} \\
&= \|f\|_p^r [\mu(X)]^{\frac{1}{s}}.
\end{aligned}$$

Hence,

$$\|f\|_r \leq \|f\|_p [\mu(X)] \frac{1}{sr} = \|f\|_p [\mu(X)]^{\frac{1}{r} - \frac{1}{p}} < \infty,$$

and so $f \in \mathcal{L}^r(X, \mu)$.

$\square$

3.5.8 Theorem (Minkowski's inequality)
Let $1 \leq 0 \leq \infty$. If $f, g \in \mathcal{L}^p(X,\mu)$, then $f+g \in \mathcal{L}^p(X,\mu)$ and

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

PROOF.

We have shown that $f+g \in \mathcal{L}^p(X,\mu)$. Let $q = \frac{p}{p-1}$. If $\int_X |f+g|^p d\mu = 0$, then the result follows trivially. Assume then that $\in_X |f+g|^p d\mu \neq 0$. Then

$$\begin{aligned}
\int_X |f+g|^p d\mu &= \int_X |f+g||f+g|^{p-1} d\mu \\
&\leq \int_X |f||f+g|^{p-1} s\mu + \int_X |g||f+g|^{p-1} d\mu \\
&\leq \Big(\int_X |f+g|^{(p-1)q}\Big)^{\frac{1}{q}} \Big[\Big(\int_X |f|^p d\mu\Big)^{\frac{1}{p}} + \Big(\int_X |g|^p d\mu\Big)^{\frac{1}{p}}\Big] \\
&= \Big(\int_X |f+g|^p d\mu\Big)^{\frac{1}{q}} [\|f\|_p + \|g\|_p].
\end{aligned}$$

Dividing both sides by $\left(\int_X |f+g|^p d\mu\right)^{\frac{1}{q}}$, gives the desired result, that is,

$$\Big(\int_X |f+g|^p d\mu\Big)^{\frac{1}{p}} = \Big(\int_X |f+g|^p d\mu\Big)^{1-\frac{1}{q}} \leq \|f\|_p + \|g\|_q \Leftrightarrow \|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

If $p = \infty$, then for each $x \in X$,

$$|f(x)+g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty.$$

Thus, $f+g \in \mathcal{L}^\infty(X,\mu)$ and

$$\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty.$$

□

Let (X, Σ, μ) be a measure space and $1 \leq p < \infty$. Define a real-valued function $\|\cdot\|_p$ on $\mathcal{L}^p(X,\mu)$ by

$$\|f\|_p := \Big(\int_X |f|^p d\mu\Big),$$

for $f \in \mathcal{L}^p(X,\mu)$.

Note that if $f \in \mathcal{L}^p(X,\mu)$, then

$$\|f\|_p = 0 \Leftrightarrow \int_X |f|^p d\mu = 0 \Leftrightarrow |f| = 0 \text{ a.e. on } X \Leftrightarrow f = 0 \text{ a.e. on } X.$$

This shows that $\|\cdot\|_p$ does **not** define a norm on $\mathcal{L}^p(X,\mu)$.

The way that we deal with this, is we define a relation $\sim$ on $\mathcal{L}^p(X,\mu)$ as follows: For $f, g \in \mathcal{L}^p(X,\mu)$,

$$f \sim g \Leftrightarrow f = g \text{ a.e. on } X.$$

It is easy to show that $\sim$ defines an equivalence relation on $\mathcal{L}^p(X,\mu)$. The relation $\sim$ identifies elements of $\mathcal{L}^p(X,\mu)$ that only differ on a null set. We denote by $L^p(X,\mu)$ the quotient space $\mathcal{L}^p(X,\mu)/\sim$. That is

$$L^p(X,\mu) = \mathcal{L}^p(X,\mu)/\mathcal{N},$$

where $\mathcal{N} = \{f \in \mathcal{L}^p(X,\mu) : f = 0 \text{ a.e. on } X\}$. Note that $L^p(X,\mu)$ is a set of equivalence classes of functions. For $f \in \mathcal{L}^p(X,\mu)$, denote by $[f]$ the equivalence class of f. That is, $[f]$ is the set of all functions in $\mathcal{L}^p(X,\mu)$ that differ from f on a null set.

Similarly, denote $L^\infty(X,\mu)$, the quotient space obtained by identifying functions in $\mathcal{L}^\infty(X,\mu)$ that coincide almost everywhere.

3.5.9 Remark
We can consider the elements of $L^p(X,\mu)$ as functions. This is technically incorrect since $L^p(X,\mu)$ is really the set of *equivalence classes of functions*. Never-the-less, we will continue to abuse the language by not distinguishing between elements $[f]$ of $L^p(X,\mu)$ and the function f representing the equivalence class $[f]$.

3.5.10 Theorem
Let (X,Σ,μ) be a measure space and $1 \le p \le \infty$. Define addition and scalar multiplication in $L^p(X,\mu)$ as follows:

$$[f]+[g] = [f+g] \text{ and } \alpha[f] = [\alpha f], \text{ where } f,g \in \mathcal{L}^p(X,\mu) \text{ and } \alpha \in \mathbb{R}.$$

Then $L^p(X,\mu)$ is a real vector space.

PROOF.

Exercise.

$\square$

Now that we know that $L^p(X,\mu)$ is a linear space, we want to define a norm on it. If $1 \le p < \infty$ and $f \in \mathcal{L}^p(X,\mu)$, define $\|\cdot\|_p : L^p(X,\mu) \to [0,\infty)$ by

$$\|[f]\|_p = \|f\|_p = \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}},$$

and for $p=\infty$, define

$$\|[f]\|_p = \|f\|_\infty = \text{esssup}_X |f|.$$

Note that it is important that the definition of $\|\cdot\|_p$ (or $\|\cdot\|_\infty$) is independent of the choice of a representative of the equivalence class $[f]$, for $f \in \mathcal{L}^p(X,\mu)$ (or $f \in \mathcal{L}^p(X,\mu)$). Indeed, let $f,g \in \mathcal{L}^p(X,\mu)$ such that $g \in [f]$. Then $[f]=[g]$ and $f=g$ a.e. on X. Therefore,

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

3.5.11 Theorem
Let (X,Σ,μ) be a measure space. Then $(L^p(X,\mu), \|\cdot\|_p)$ and $(L^\infty(X,\mu), \|\cdot\|_\infty)$ are normed linear spaces.

3.5.12 Example
Let $X=\mathbb{N}, \Sigma = \mathcal{P}(\mathbb{N})$, and ν be the counting measure on $(\mathbb{N},\mathcal{P}(\mathbb{N}))$. Then the functions on $\mathbb{N}$ are sequences, integration is summation, and integrable functions correspond to absolutely summable sequences.

For $1 \leq p < \infty$, then $L^p(\mathbb{N}, \nu) = l^p(\mathbb{N})$, the spaces of real sequences (x_n) such that $\sum_{n\in\mathbb{N}} |x_n|^p < \infty$. The norm on $l^p(\mathbb{N})$ is

$$\|(x_n)\|_p = \left(\sum_{n\in\mathbb{N}} |x_n|^p\right)^{\frac{1}{p}}.$$

If $p = \infty$, then $L^\infty(\mathbb{N}\nu) = l^\infty(\mathbb{N})$, then the spaces of bounded real sequences (x_n) and

$$\|(x_n)\|_\infty = \sup_{n\in\mathbb{N}} |x_n|.$$

Chapter 4

Decomposition of measures

4.1 Signed measures

In this chapter, we discuss the Hahn decomposition, Jordan decomposition, Radon-Nikodym Theorem and the Lebesgue decomposition Theorem, which are amongst the most important results in measure theory.

4.1.1 Definition
Let (X, Σ) be a measurable space. A **signed measure** on Σ is an extended real-valued set function ν such that

(a) ν assumes at most one of the values ∞ and $-\infty$,

(b) $\nu(\emptyset) = 0$,

(c) for any sequence (A_n) of disjoint measurable sets,

$$\nu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \nu(A_n).$$

It follows from the above definition that every measure space is a signed measure, but not conversely.

4.1.2 Example
Let (X, Σ, μ) be a measure space and f a measurable function on X such that f^+ or f^- is integrable. Define ν on Σ by

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

for $A \in \Sigma$. Then ν is a signed measure on Σ.

4.1.3 Definition
Let ν be a signed measure on a measurable space (X, Σ). A set $A \in \Sigma$ is said to be a

[1] **positive set** with respect to ν if $\nu(A \cap E) \geq 0$, for any $E \in \Sigma$.

[2] **negative set** with respect to ν if $\nu(A \cap E) \leq 0$, for any $E \in \Sigma$.

4.1.4 Remarks
[1] It can be shown tht a set $A \in \Sigma$ is positive (resp., negative) with respect to ν if and only if $\nu(E) \geq 0$ (resp., $\nu(E) \leq 0$) for every measurable subset E of A.

[2] One can also show that $A \in \Sigma$ is a null set if and only if $\nu(E) = 0$, for all $E \in \Sigma$ such that $E \subset A$.

[3] Any measurable subset of a positive (resp., negative) set is positive (resp., negative). Also, a countable union of positive (resp., negative) sets is positive (resp., negative).

We show the latter for positive sets: Let (A_n) be a sequence of positive sets in Σ. then there is a sequence (B_n) in Σ such that $B_n \subset A_n$, for each $n \in \mathbb{N}$, $B_i \cap B_j = \emptyset$ if $i \neq j$ and $\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} B_n$.

Let $E \in \Sigma$ such that $E \subset \bigcup_{n\in\mathbb{N}} A_n$. Then, since

$$E = E \cap \left(\bigcup_{n\in\mathbb{N}} A_n\right) = E \cap \left(\bigcup_{n\in\mathbb{N}} B_n\right) = \bigcup_{n\in\mathbb{N}}(E \cap B_n) \text{ (a disjoint set)},$$

it follows that

$$\nu(E) = \sum_{n=1}^{\infty} \nu(E \cap B_n).$$

Since A_n is positive for each $n \in \mathbb{N}$ and $E \cap B_n \subset A_n$, it follows that $\nu(E \cap B_n) \geq 0$, for each $n \in \mathbb{N}$. Hence, $\nu(E) \geq 0$. That is, $\bigcup_{n\in\mathbb{N}} A_n$ is a positive set.

[4] If A is a positive for each set with respect to ν, then $\nu(A) = 0$. The converse is not true in general.

4.1.5 Theorem

Let (X, Σ) be a measurable space and ν a signed measure on Σ. If $E \in \Sigma$ and $0 < \nu(E) < \infty$, then E contains a positive set A with $\nu(A) > 0$.

PROOF.

If E is positive with respect to ν, then take $A = E$. Assume that E is not positive. Then there is a set $B \in \Sigma$ with $B \subset E$ and $\nu(B) < 0$. Let k_1 be the smallest positive integer such that there is an $E_1 \in \Sigma$ with $E_1 \subset E$ and $\nu(E_1) < -\frac{1}{k_1}$. Since

$$0 < \nu(E) + \frac{1}{k_1} < \nu(E) - \nu(E_1) = \nu(E \backslash E_1) < \infty,$$

we can re[eat the above argument used for E in the case of $E \backslash E_1$. Let k_1 be the smallest positive integer such that there is an $E_2 \in \Sigma$ with $E_2 \subset E \backslash E_1$ and $\nu(E_2) < -\frac{1}{k_2}$. Then

$$0 < \nu(E \backslash E_1) + \frac{1}{k_2} < \nu(E \backslash E_1) - \nu(E_2) = \nu(E \backslash (E_1 \cup E_2)) < \infty.$$

Continuing in this fashion, we obtain a sequence (k_n) or positive integers such that for each $n \in \mathbb{N}$, k_n is the smallest positive integer for which there is a measurable set E_n such that

$$E_n \subset E \backslash \bigcup_{j=1}^{n-1} E_j$$

and $\nu(E_n) \subset E \backslash \bigcup_{j=1}^{n-1} E_j$ and $\nu(E_n) < -\frac{1}{k_n}$. If the process stops at n_s, then let $A = E \backslash \bigcup_{j=1}^{s} E_j$. Then A is a positive set and $\nu(A) > 0$ (if $\nu(A) = 0$, then $\nu(E) = \sum_{j=1}^{s} \nu(E_j) < 0$, contradicting the fact that $\nu(E) > 0$). Otherwise, let $A = E \backslash \bigcup_{j=1}^{\infty} E_j$. Then $E = A \cup \bigcup_{j=1}^{\infty} \nu(E_j)$, as disjoint union. Hence,

$$\nu(E) = \nu(A) + \sum_{j=1}^{\infty} \nu(E_j).$$

If $\nu(A) = 0$, then $\nu(E) = \sum_{j=1}^{\infty} \nu(E_j) < 0$, contradicting the hypothesis that $\nu(E) > 0$. Similarly, if $\nu(A) < 0$, then $\nu(E) < 0$. Thus $\nu(A) > 0$. Since $0 < \nu(E) < \infty$, the series $\sum\limits_{j=1}^{\infty} \nu(E_j)$ converges. Therefore $-\infty < \sum\limits_{j=1}^{\infty} \nu(E_j)$, whence $\sum\limits_{j=1}^{\infty} \frac{1}{k_j} < \infty$. In particular, we have that $k_j \to \infty$ as $j \to \infty$. We show that A is a positive set. Since $\lim_{n\to\infty} k_j = \infty$, there is a $j_0 \in \mathbb{N}$ such that $k_j > 1$, for all $j \geq j_0$. Let B be a measurable subset of A and $j > j_0$. Then $B \subset E \backslash \bigcup\limits_{i=1}^{j} E_i$ and so

$$\nu(B) \geq -\frac{1}{k_j - 1}.$$

This is true for all $j \geq j_0$. Letting $j \to \infty$, we get that $\nu(B) \geq 0$. Hence A is a positive set.

□

4.1.6 Theorem (Hahn decomposition)
Let ν be a signed measure on a measurable space (X, Σ). Then there is a pair (P, N) of measurable subsets of X such that

[1] P is a positive set and N a negative set.

[2] $P \cap N = \emptyset$.

[3] $X = P \cup N$.

PROOF.

Since ν assumes at most one of the values of $-\infty$ or ∞, we may assume that ∞ is the value that is never attained. That is, $\nu(E) < \infty$, for all $E \in \Sigma$. Let $\wp$ be the collection of all sets in Σ that are positive with respect to ν. Then $\wp$ is nonempty since $\emptyset \in \wp$. Let

$$\alpha = \sup\{\nu(A) : A \in \wp\}.$$

Then $\alpha \geq 0$ since $\emptyset \in \wp$. Let (A_n) be a sequence in $\wp$ such that $\alpha = \lim_{n\to\infty} \nu(A_n)$ and let $P = \bigcup_{n\in\mathbb{N}} A_n$. Since a countable union of positive sets is positive, we have that P is positive with respect to ν. By the definition of α, we have that $\alpha \geq \nu(P)$. Since $P\backslash A_n \subset P$, for each $n \in \mathbb{N}$ and P is positive with respect to ν, it follows that $\nu(P\backslash A_n) \geq 0$, for each $n \in \mathbb{N}$. Therefore,

$$\nu(P) = \nu(A_n) + \nu(P\backslash A_n) \geq \nu(A_n),$$

for each $n \in \mathbb{N}$. Thus, $\alpha \leq \nu(P)$. Hence, $\nu(P) = \alpha$.

Let $N = X\backslash P$. We show that N is negative with respect to ν. Let E be a measurable subset of N. If $\nu(E) > 0$, then, by Theorem 4.1.5, E contains a subset B such that $\nu(B) > 0$. But then $P \cup B$ (a disjoint union) would be a positive set with $\nu(P \cup B) = \nu(P) + \nu(B) > \alpha$, contradicting the definition of α.

It is clear from the above that $P \cap N = \emptyset$ and $X = P \cup N$.

□

4.1.7 Definition
The pair of sets (P, N) as discussed in Theorem 4.1.6, is called a **Hahn decomposition** of X with respect to the signed measure ν.

The Hahn decomposition (P, N) is not unique. Indeed, if M is a null set with respect to ν, then the pairs $(P \cup M, N\backslash M)$ and $(P\backslash M, N \cup M)$ are also Hahn decompositions of X with respect to ν. Despite this lack of uniqueness, we have the following proposition.

4.1.8 Proposition
Let (X, Σ) be a measurable space and ν a signed measure on X. If (P_1, N_1) and (P_2, N_2) are Hahn decompositions of X with respect to ν, then, for all $E \in \Sigma$,

$$\nu(E \cap P_1) = \nu(E \cap P_2) \text{ and } \nu(E \cap N_1) = \mu(E \cap N_2).$$

PROOF.

We firstly note that since

$$P_1 \backslash P_2 \subset P_1 \text{ and } P_1 \backslash P_2 = P_1 \cap P_2^c = P_1 \cap N_2,$$

we have that, for any $E \in \Sigma$,

$$E \cap (P_1 \backslash P_2) \subset E \cap P_1 \cap P_1 \text{ and } E \cap (P_1 \backslash P_2) = E \cap (P_1 \cap N_2) \subset N_2.$$

Therefore,

$$\nu(E \cap (P_1 \backslash P_2)) \geq 0 \text{ and } \nu(E \cap (P_1 \backslash P_2)) \leq 0.$$

It now follows that $\nu(E \cap (P_1 \backslash P_2)) = 0$. A similar argument shows that $\nu(E \cap (P_2 \backslash P_1)) = 0$.

Now,

$$\begin{aligned} E \cap P_1 &= E \cap [(P_1 \backslash P_2) \cup (P_1 \cap P_2)] \\ &= (E \cap (P_1 \backslash P_2)) \cup (E \cap (P_1 \cap P_2)) \text{ (a disjoint union).} \end{aligned}$$

Thus,

$$\nu(E \cap P_1) = \nu(E \cap (P_1 \backslash P_2)) + \nu(E \cap (P_1 \cap P_2)) = \nu(E \cap (P_1 \cap P_2)).$$

Similarly,

$$\begin{aligned} E \cap P_2 &= E \cap [(P_2 \backslash P_1) \cup (P_1 \cap P_2)] \\ &= (E \cap (P_2 \backslash P_1)) \cup (E \cap (P_1 \cap P_2)) \text{ (a disjoint union).} \end{aligned}$$

Therefore

$$\nu(E \cap P_2) = \nu(E \cap (P_2 \backslash P_1)) + \nu(E \cap (P_1 \cap P_2)) = \nu(E \cap (P_1 \cap P_2)).$$

It follows from this that

$$\nu(E \cap P_1) = \nu(E \cap P_2).$$

The proof that $\nu(E \cap N_1) = \nu(E \cap N_2)$ is left as an exercise to the reader.

□

4.1.9 Definition
Let ν be a signed measure on a measurable space (X, Σ) and let $\{P, N\}$ be a Hahn decomposition of X with respect to ν. The **positive variation** ν^+, **negative variation** ν^-, and **total variation** $|\nu|$ of ν are finite measures defined on Σ by

$$\begin{aligned} \nu^+(E) &= \nu(E \cap P) \\ \nu^-(E) &= -\nu(E \cap N) \\ |\nu|(E) &= \nu(E \cap P) - \nu(E \cap N) = \nu^+(E) + \nu^-(E). \end{aligned}$$

It is clear from Proposition 4.1.8, that the definitions of ν^+, ν^-, and $|\nu|$ are independent of any Hahn decomposition of X with respect to the signed measure ν.

Let (P, N) be a Hahn decomposition of X with respect to a signed measure ν. Then $X = P \cup N$ and $P \cap N = \emptyset$. Let $E \in \Sigma$, then

$$E = E \cap X = E \cap (P \cup N) = (E \cap P) \cup (E \cap N), \text{ a disjoint union.}$$

Therefore,

$$\nu(E) = \nu(E \cap P) + \nu(E \cap N) = \nu^+(E) - \nu^-(E).$$

Hence,

$$\nu = \nu^+ - \nu^-. \text{ (This is called the \textbf{Jordan decompostion of} } \nu.)$$

Since ν assumes at moest one of the values $-\infty$ and ∞, either ν^+ or ν^- must be finite. If both ν^+ and ν^- are finite, then ν is called a finite signed measure.

4.1.10 Example

Let (X, Σ, μ) be a measure space and $f \in L^1(\mu)$. Let ν be a set function defined on Σ by

$$\nu(E) = \int_E f d\mu, E \in \Sigma.$$

Show that ν is a signed measure on X and

$$\nu^+(E) = \int_E f^+ d\mu, \nu^-(E) = \int_E f^- d\mu, \text{ and } |\nu|(E) = \int_E |f| d\mu.$$

Solution: That ν is a signed measure on X follows from Proposition 3.3.15.

Let $P = \{x \in X : f(x) > 0\}$ and $P = \{x \in X : f(x) \leq 0\}$. Then $X = P \cup N$ and $P \cap N = \emptyset$. Let $E \in \Sigma$. Then

$$\begin{aligned} \nu(E \cap P) &= \int_{E \cap P} f d\mu = \int_{\{x \in E : f(x) > 0\}} f d\mu \\ &= \int_E f^+ d\mu \geq 0, \end{aligned}$$

and

$$\begin{aligned} \nu(E \cap N) &= \int_{E \cap N} f d\mu = \int_{\{x \in X : f(x) \leq 0\}} f d\mu \\ &= -\int_E f^- d\mu \leq 0. \end{aligned}$$

Thus, P is a positive set with respect to ν and N is a negative set with respect to ν. that is, (P, N) is a Hahn decomposition of X with respect to ν. By definition of ν^+, ν^-, and $|\nu|$ we have that, for each $E \in \Sigma$,

$$\begin{aligned} \nu^+(E) &= \nu(E \cap P) = \int_E f^+ d\mu, \\ \nu^-(E) &= -\nu(E \cap N) = \int_E f^- d\mu, \\ |\nu|(E) &= \nu^+(E) + \nu^-(E) = \int_E f^+ d\mu + \int_E f^- d\mu = \int_E (f^+ + f^-) d\mu = \int_E |f| d\mu. \end{aligned}$$

4.1.11 Definition

Two measures λ and μ on a measurable space (X, Σ) are said to be **mutually singular**, denoted by $\lambda \perp \mu$, if there is an $A \in \Sigma$ such that $\lambda(A) = 0 = \mu(A^c)$.

If λ and μ are signed measures, then $\lambda \perp \mu$ if $|\lambda| \perp |\mu|$.

4.1.12 Example
Let ν be a signed measure on a measurable space (X, Σ). Then the measures ν^+ and ν^- are mutually singular. Indeed, let (P, N) be a Hahn decomposition of X with respect to ν. Then $X = P \cup N$ and $P \cap N = \emptyset$. Therefore,

$$\nu^+(N) = \nu(N \cap P) = \nu(\emptyset) = 0 \text{ and}$$
$$\nu^-(P) = -\nu(N \cap P) = -\nu(\emptyset) = 0.$$

Thus $\nu^+ \perp \nu^-$.

4.1.13 Definition
Let λ and ν be measures on a measurable space (X, σ). We say that λ is **absolutely continuous with respect to** μ, denoted by $\lambda \ll \mu$, if $\lambda(E) = 0$, for every $E \in \Sigma$ for which $\mu(E) = 0$.

If μ and λ are signed measures, then $\lambda \ll \mu$ if $|\lambda| \ll |\mu|$.

4.1.14 Example
Let (X, Σ, μ) be a measure space and f a nonnegative measurable function on X. We have shown in Proposition 3.3.15 that the set function

$$\lambda(E) = \int_E f d\mu, \ E \in \Sigma \tag{4.1}$$

defines a measure on X. The measure λ will be finite if and only if f is integrable. Since the integral over a set $E \in \Sigma$, with $\mu(E) = 0$ is zero, it follows that λ is absolutely continuous with respect to μ.

4.1.15 Proposition
Let λ, μ, and ν be measures on a measurable space (X, Σ) and c_1, c_2 numbers. Then

[1] If $\lambda \ll \mu$ and $\nu \ll \mu$, then $(c_1\lambda + c_2\nu) \ll \mu$.

[2] If $\lambda \perp \mu$ and $\nu \perp \mu$, then $(c_1\lambda + c_2\nu) \perp \mu$.

[3] If $\lambda \perp \mu$ and $\lambda \perp \mu$, then $\lambda = 0$.

PROOF.

We prove [3] and leave [1] and [2] as exercises.

Since $\lambda \perp \mu$, there is a set $E \in \Sigma$ such that $\lambda(E) = 0 = \mu(E^c)$. Using the fact that $\lambda \ll \mu$, it follows that $\lambda(A^c) = 0$. Therefore,

$$\lambda(X) = \lambda(A \cup A^c) = \lambda(A) + \lambda(A^c) = 0.$$

Hence, $\lambda = 0$.

□

The following theorem, known as the Radon-Nikodym Theorem shows that if (X, Σ, μ) is a σ-finite measure space, then every measure on X that is absolutely continuous with respect to μ arises in the manner indicated by (4.1).

4.1.16 Theorem (Radon-Nikodym Theorem)
Let ν and μ be σ-finite measures on a measurable space (X, Σ). Suppose that ν is absolutely continuous with respect to μ. Then there is a nonnegative measurable function f on X such that, for each $E \in \Sigma$,

$$\nu(E) = \int_E f d\mu.$$

Moreover, the function f is uniquely determined almost everywhere $[\mu]$.

The function f as described in the Radon-Nikodym Theorem is also called the **Radon-Nikodym derivative of** ν with respect to μ. We write $f = \frac{d\nu}{d\mu}$ or $d\nu = f d\mu$.

4.1.17 Theorem (Lebesgue Decomposition Theorem)
Let (X, Σ, μ) be a σ-finite measure space and ν a σ-finite measure on Σ. Then there is a unique pair of measures ν_a and ν_s, with $\nu = \nu_a + \nu_s$ such that $\nu_a \ll \mu$ and $\nu_s \perp \mu$.

PROOF.

Let $\lambda = \mu + \nu$. Then λ is a σ-finite measure with $\mu \ll \lambda$ and $\nu \ll \lambda$. By the Radon-Nikodym Theorem, there are nonnegative measurable functions f and g on X such that

$$\mu(E) = \int_E f d\lambda \text{ and } \nu(E) = \int_E g d\lambda, \ E \in \Sigma.$$

Let $A = \{x \in X : f(x) > 0\}$ and $B = \{x \in X : f(x) = 0\}$. Then $X = A \cup B$, $A \cap B = \emptyset$, and $\mu(B) = \int_B f d\lambda = 0$. Define ν_a and ν_s by

$$\nu_a(E) = \nu(E \cap A) \text{ and } \nu_s(E) = \nu(E \cap B).$$

Let $E \in \Sigma$. Then

$$E = E \cap X = E \cap (A \cup B) = (E \cap A) \cup (E \cap B), \text{ a disjoint union.}$$

Therefore,

$$\nu(E) = \nu(E \cap A) + \nu(E \cap B) = \nu_a(E) + \nu_s(E) = (\nu_a + \nu_s)(E).$$

Hence, $\nu = \nu_a + \nu_s$.

Since, $\nu_s(A) = \nu(A \cap B) = \nu(\emptyset) = 0$, we have that $\nu_s(A) = 0 = \mu(B)$. That is, $\nu_s \perp \mu$. We now show that $\nu_a \ll \mu$. Let $E \in \Sigma$ such that $\mu(E) = 0$. Then

$$0 = \mu(E) = \int_E f d\lambda.$$

Therefore, $f = 0$ a.e. $[\lambda]$. Now,

$$0 \leq \int_{A \cap E} f d\lambda \leq \int_E f d\lambda = 0.$$

Hence, by Proposition 3.3.7, $\int_{A \cap E} f d\lambda = 0$. Since $f > 0$ on $A \cap E$, we must have that $\lambda(A \cap E) = 0$. Since $\nu \ll \lambda$, it follows that $\nu(A \cap E) = 0$ and therefore $\nu_a(E) = 0$. That is, $\nu_a \ll \mu$.

For uniqueness, ;et ω_a and ω_s be another pair of measures with $\nu = \omega_a + \omega_s$ such that $\omega_a \ll \mu$ and $\omega_s \perp \mu$. Then

$$\nu_a + \nu_s = \omega_a + \omega_s \Leftrightarrow \nu_a - \omega_a = \omega_s - \nu_s.$$

Hence, $(\nu_a - \omega_a) \ll \mu$ and $(\omega_s - \nu_s) \perp \mu$. But since $\nu_a - \omega_a = \omega_s - \nu_s$, it follows that $(\nu_a - \omega_a) \perp \mu$. That is,

$$(\nu_a - \omega_a) \ll \mu, \quad (\nu_a - \omega_a) \perp \mu \text{ and}$$
$$(\omega_s - \nu_s) \ll \mu \quad (\omega_s - \nu_s) \perp \mu.$$

By Proposition 4.1.15, we have that $\nu_a - \omega_s = 0$ and $\omega_s - \nu_s = 0$. Hence, $\nu_a = \omega_a$ and $\nu_s = \omega_s$.

$\square$