

# Measure Theory and Fourier Analysis

Summary Notes for MATH3969

Shirley Wang

Semester 2, 2025

## Contents

1	$\sigma$ -algebras	3
2	Measures	4
3	The construction of measures from outer measures	5
4	The Lebesgue measure	7
5	Regularity of the Lebesgue Measure	7
6	Uniqueness of the Lebesgue Measure	8
7	Lebesgue Measure and Linear Transformations	8
8	The Lebesgue-Stieltjes measure	9
9	Measurable functions	10
10	Sequences of measurable functions	11
11	Simple measurable functions	11
12	Approximation by simple functions	11
13	The integration of non-negative simple functions	12
14	Integration of non-negative measurable functions	12
15	The monotone convergence theorem	13
16	Integrable functions	13
17	Limit Theorems	14
18	The $\mathcal{L}^p$ -spaces	16
19	The $L^p$ -spaces	17

20 Basic density theorems	18
21 The space of bounded measurable functions	18
22 Fubini's Theorem	18
23 Translation of a function	19
24 Convex functions and Jensen's inequality	20
25 Convolution	20
26 Approximate identities	21
27 Approximation theorems	21
28 Definition and basic properties of the Fourier transform	22
29 Fundamental properties of the Fourier transform	22
30 The Fourier transform on $L^2(\mathbb{R}^N)$	23
31 The Riesz representation theorem	23
32 The Radon-Nikodym theorem	24
33 Basic concepts of probability theory	25
34 The distribution of a random variable	25
35 Conditional expectation	26

## Introduction

**Definition 0.1** (0.1). *Let  $X$  be a set and  $\mathcal{A}$  a collection of subsets of  $X$ . We call  $\mathcal{A}$  a  $\sigma$ -algebra if*

- (i)  $\emptyset \in \mathcal{A}$ ;
- (ii)  $A \in \mathcal{A}$  implies  $A^c \in \mathcal{A}$ ;
- (iii)  $A_j \in \mathcal{A}$  for all  $j \in \mathbb{N}$  implies  $\bigcup_{j \in \mathbb{N}} A_j \in \mathcal{A}$ .

**Definition 0.2** (0.2). *Let  $X$  be a set and  $\mathcal{A}$  a  $\sigma$ -algebra of subsets of  $X$ . A function  $\mu: \mathcal{A} \rightarrow [0, \infty]$  is called a measure if*

- (i)  $\mu(\emptyset) = 0$ ;
- (ii)  $A_j \in \mathcal{A}$ ,  $j \in \mathbb{N}$  are such that  $A_k \cap A_j = \emptyset$  for  $j \neq k$ , then  $\mu\left(\bigcup_{j \in \mathbb{N}} A_j\right) = \sum_{j=0}^{\infty} \mu(A_j)$ .

# 1 $\sigma$ -algebras

**Definition 1.1** (1.1). If  $X$  is a set we call the class of all subsets of  $X$  the power set of  $X$  and denote it by  $\mathcal{P}(X)$ .

**Definition 1.2** (1.2). Let  $\mathcal{A}$  be a collection of subsets of a set  $X$ , that is,  $\mathcal{A} \subseteq \mathcal{P}(X)$ . We call  $\mathcal{A}$  a  $\sigma$ -algebra of subsets of  $X$  if it has the following properties:

- (i)  $\emptyset \in \mathcal{A}$ ;
- (ii) If  $A \in \mathcal{A}$ , then  $A^c \in \mathcal{A}$ ;
- (iii) If  $A_k \in \mathcal{A}$  for all  $k \in \mathbb{N}$ , then  $\bigcup_{k \in \mathbb{N}} A_k \in \mathcal{A}$ .

We call  $\mathcal{A}$  an algebra if instead of (iii) we only have

- (iii) If  $A, B \in \mathcal{A}$ , then  $A \cup B \in \mathcal{A}$ .

**Proposition 1.1** (1.5). Let  $\mathcal{A}$  be a  $\sigma$ -algebra of subsets of  $X$ . Then the following statements are true:

- (i)  $\emptyset, X \in \mathcal{A}$ ;
- (ii) If  $A_k \in \mathcal{A}$ ,  $k \in \mathbb{N}$ , then  $\bigcap_{k \in \mathbb{N}} A_k \in \mathcal{A}$ ;
- (iii) If  $A, B \in \mathcal{A}$ , then  $A \cup B \in \mathcal{A}$  and  $A \cap B \in \mathcal{A}$ .

**Proposition 1.2** (1.6). Let  $I$  be an arbitrary index set and suppose that for every  $i \in I$ ,  $\mathcal{A}_i$  is a  $\sigma$ -algebra of subsets of  $X$ . Then

$$\mathcal{A} := \bigcap_{i \in I} \mathcal{A}_i \subseteq \mathcal{P}(X)$$

is a  $\sigma$ -algebra of subsets of  $X$ .

**Definition 1.3** (1.7). (a) Let  $C$  be a collection of subsets of  $X$ . We call  $\mathcal{A}(C)$  the  $\sigma$ -algebra generated by  $C$ .

(b) Let  $X$  be a metric (or more generally a topological) space and  $C$  the collection of all open subsets of  $X$ . Then  $\mathcal{B} := \mathcal{A}(C)$  is called the Borel  $\sigma$ -algebra in  $X$ . Sets in  $\mathcal{B}$  are called Borel sets.

**Definition 1.4** (1.9). Let  $X$  be a set and  $A_k$ ,  $k \in I$ , be a collection of subsets, where  $I$  is an arbitrary index set. We say that this collection consists of disjoint sets if  $A_j \cap A_k = \emptyset$  whenever  $j \neq k$ .

**Lemma 1.1** (1.10). Let  $A_n \in \mathcal{X}$  for  $n \in \mathbb{N}$ . Set

$$B_0 := A_0 \quad \text{and} \quad B_n := A_n \cap (A_0 \cup A_1 \cup \cdots \cup A_{n-1})^c \quad \text{for } n \geq 1.$$

Then  $\bigcup_{k=0}^n B_k = \bigcup_{k=0}^n A_k$  for all  $n \geq 0$  and  $B_k$ ,  $k \in \mathbb{N}$ , is disjoint.

**Proposition 1.3** (1.11). Suppose that  $\mathcal{A}$  is a collection of subsets of  $X$  with the following properties.

- (i)  $\emptyset \in \mathcal{A}$ ;

- (ii) If  $A \in \mathcal{A}$ , then  $A^c \in \mathcal{A}$ ;
- (iii) If  $A_k \in \mathcal{A}$ ,  $k \in \mathbb{N}$ , are disjoint, then  $\bigcup_{k \in \mathbb{N}} A_k \in \mathcal{A}$ ;
- (iv) If  $A, B \in \mathcal{A}$ , then  $A \cap B \in \mathcal{A}$ .

Then  $\mathcal{A}$  is a  $\sigma$ -algebra.

**Proposition 1.4** (1.12). Suppose that  $X, Y$  are sets and  $f: X \rightarrow Y$  a function.

- (i) If  $\mathcal{A}$  is a  $\sigma$ -algebra in  $Y$ , then

$$\mathcal{A}_0 := \{f^{-1}[A] : A \in \mathcal{A}\}$$

is a  $\sigma$ -algebra in  $X$ .

- (ii) If  $\mathcal{A}$  is a  $\sigma$ -algebra in  $X$ , then

$$\mathcal{A}_1 := \{A \subseteq Y : f^{-1}[A] \in \mathcal{A}\}$$

is a  $\sigma$ -algebra in  $Y$ .

**Definition 1.5** (1.13). Let  $f: X \rightarrow Y$  be a function and  $\mathcal{A}$  a  $\sigma$ -algebra of subsets of  $Y$ . The  $\sigma$ -algebra  $\mathcal{A}_0 := \{f^{-1}[A] : A \in \mathcal{A}\}$  is called the  $\sigma$ -algebra generated by  $f$ .

## 2 Measures

**Definition 2.1** (2.1). Let  $\mathcal{A}$  be a  $\sigma$ -algebra of subsets of  $X$ . A function  $\mu: \mathcal{A} \rightarrow [0, \infty]$  is called a measure if

- (i)  $\mu(\emptyset) = 0$ ;
- (ii) for every countable collection  $A_k \in \mathcal{A}$ ,  $k \in \mathbb{N}$ , of disjoint sets

$$\mu\left(\bigcup_{k=0}^{\infty} A_k\right) = \sum_{k=0}^{\infty} \mu(A_k) \quad (\text{countable additivity}).$$

Moreover, if  $A \in \mathcal{A}$  we call the set  $A$  measurable (with respect to  $\mu$ ). Finally the triple  $(X, \mathcal{A}, \mu)$  is called a measure space. If  $\mu(X) = 1$  we sometimes call  $\mu$  a probability measure and  $(X, \mathcal{A}, \mu)$  a probability space. If  $X$  is a metric space, we call  $\mu$  a Borel measure if  $\mathcal{B} \subseteq \mathcal{A}$ , that is, if all Borel sets are  $\mu$ -measurable.

**Example 2.1** (2.2). Let  $X$  be a set and fix  $x \in X$ . For  $A \subseteq X$  we set

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

Then  $\delta_x: \mathcal{P}(X) \rightarrow [0, \infty)$  is a measure. It is called the Dirac measure concentrated at  $x$ .

**Example 2.2** (2.3). Let  $X$  be a set and fix  $x \in X$ . For  $A \subseteq X$  we set

$$\mu(A) := \begin{cases} \#A & \text{if } A \text{ has finite cardinality,} \\ \infty & \text{otherwise,} \end{cases}$$

where  $\#A$  is the cardinality of  $A$ . Then  $\mu: \mathcal{P}(X) \rightarrow [0, \infty]$  is a measure. It is called the counting measure on  $X$ .

**Proposition 2.1** (2.5). Let  $(X, \mathcal{A}, \mu)$  be a measure space,  $Y$  a set and  $f: X \rightarrow Y$  a function. Define  $\mathcal{A}_1 = \{A \subseteq Y: f^{-1}[A] \in \mathcal{A}\}$  and

$$\mu_f(A) := \mu(f^{-1}[A])$$

for all  $A \in \mathcal{A}_1$ . Then  $(Y, \mathcal{A}_1, \mu_f)$  is a measure space.

**Proposition 2.2** (2.6). Let  $(X, \mathcal{A}, \mu)$  be a measure space.

- (i) If  $A, B \in \mathcal{A}$  and  $A \cap B = \emptyset$ , then  $\mu(A \cup B) = \mu(A) + \mu(B)$ .
- (ii) If  $A, B \in \mathcal{A}$  and  $A \subseteq B$ , then  $\mu(A) \leq \mu(B)$ .
- (iii) If  $A, A_k \in \mathcal{A}$ ,  $k \in \mathbb{N}$ , and  $A \subseteq \bigcup_{k=0}^{\infty} A_k$ , then

$$\mu(A) \leq \sum_{k=0}^{\infty} \mu(A_k).$$

The last property is referred to as the countable sub-additivity of a measure.

**Proposition 2.3** (2.7). Let  $(X, \mathcal{A}, \mu)$  be a measure space.

- (i) If  $A_k \in \mathcal{A}$ ,  $k \in \mathbb{N}$ , and  $A_0 \subseteq A_1 \subseteq A_2 \subseteq \dots$ , then

$$\lim_{k \rightarrow \infty} \mu(A_k) = \mu\left(\bigcup_{k=0}^{\infty} A_k\right).$$

- (ii) If  $A_k \in \mathcal{A}$ ,  $k \in \mathbb{N}$ , with  $A_0 \supseteq A_1 \supseteq A_2 \supseteq \dots$  and  $\mu(A_k) < \infty$  for some  $k \in \mathbb{N}$ , then

$$\lim_{k \rightarrow \infty} \mu(A_k) = \mu\left(\bigcap_{k=0}^{\infty} A_k\right).$$

**Remark 2.1** (2.8). Without the assumption that  $\mu(A_k) < \infty$  (and therefore  $\mu(A_j) < \infty$  for all  $j \geq k$ ) it is possible that (ii) is not true.

### 3 The construction of measures from outer measures

**Definition 3.1** (3.1). For every subset  $A \subseteq \mathbb{R}^N$  we set

$$m_N^*(A) := \inf \left\{ \sum_{k=0}^{\infty} \text{vol}(R_k) : R_k, k \in \mathbb{N}, \text{ open rectangles with } A \subseteq \bigcup_{k=0}^{\infty} R_k \right\}.$$

We call  $m_N^*(A)$  the Lebesgue outer measure of the set  $A$ .

**Definition 3.2** (3.3). A function  $\mu^*: \mathcal{P}(X) \rightarrow [0, \infty]$  is called an outer measure if

(i)  $\mu^*(\emptyset) = 0$ ;

(ii) for every countable collection  $A, A_k \subseteq X$ ,  $k \in \mathbb{N}$  with  $A \subseteq \bigcup_{k=0}^{\infty} A_k$

$$\mu^*(A) \leq \sum_{k=0}^{\infty} \mu^*(A_k) \quad (\text{countable sub-additivity}).$$

**Definition 3.3** (3.5). Let  $\mu^*: \mathcal{P}(X) \rightarrow [0, \infty]$  be an outer measure on the set  $X$ . We call a set  $A \subseteq X$  a  $\mu^*$ -measurable set if

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

for all  $S \subseteq X$ .

**Remark 3.1** (3.6). Since  $S = (S \cap A) \cup (S \cap A^c)$  for all sets  $A, S \subseteq X$  it follows that

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

for all  $A, S \subseteq X$ . Hence to prove that  $A$  is  $\mu^*$ -measurable we only need to show that

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

for all  $S \subseteq X$ .

**Proposition 3.1** (3.7). Let  $\mu^*$  be an outer measure and  $\mathcal{A}$  the set of  $\mu^*$ -measurable sets. Then the following assertions are true.

(i) If  $\mu^*(A) = 0$ , then  $A \in \mathcal{A}$ ;

(ii) If  $A \in \mathcal{A}$ , then  $A^c \in \mathcal{A}$ ;

(iii) If  $A \in \mathcal{A}$ ,  $B, S \subseteq X$  and  $A \cap B = \emptyset$ , then

$$\mu^*(S \cap (A \cup B)) = \mu^*(S \cap A) + \mu^*(S \cap B);$$

(iv) If  $A, B \in \mathcal{A}$ , then  $A \cup B \in \mathcal{A}$  and  $A \cap B \in \mathcal{A}$ .

**Theorem 3.1** (3.8: Carathéodory). Let  $\mu^*: \mathcal{P}(X) \rightarrow [0, \infty]$  be an outer measure and let

$$\mathcal{A} := \{A \subseteq X: \mu^*(S) = \mu^*(S \cap A) + \mu^*(S \cap A^c) \text{ for all } S \subseteq X\}.$$

Then  $\mathcal{A}$  is a  $\sigma$ -algebra and  $\mu := \mu^*|_{\mathcal{A}}: \mathcal{A} \rightarrow [0, \infty]$  is a measure.

## 4 The Lebesgue measure

**Definition 4.1** (4.1). *We call*

$$\mathcal{M}_N := \{A \subseteq \mathbb{R}^N : m_N^*(S) = m_N^*(S \cap A) + m_N^*(S \cap A^c) \text{ for all } S \subseteq \mathbb{R}^N\}$$

*the Lebesgue  $\sigma$ -algebra and  $m_N := m_N^*|_{\mathcal{M}_N}$  the ( $N$ -dimensional) Lebesgue measure. Sets in  $\mathcal{M}_N$  are called Lebesgue measurable or simply measurable subsets of  $\mathbb{R}^N$ .*

**Proposition 4.1** (4.2). *Let  $R$  be an open rectangle in  $\mathbb{R}^N$ . Then*

$$m_N^*(R) = m_N^*(\bar{R}) = \text{vol}(R),$$

*where  $\bar{R}$  is the closure of  $R$ .*

**Proposition 4.2** (4.3). *For every  $\delta > 0$  and  $A \subseteq \mathbb{R}^N$  we have*

$$m_N^*(A) = \inf \left\{ \sum_{k=0}^{\infty} \text{vol}(Q_k) : Q_k \text{ open cube, } A \subseteq \bigcup_{k=0}^{\infty} Q_k, \text{diam}(Q_k) < \delta \right\}.$$

**Proposition 4.3** (4.4). *Let  $A, B \subseteq \mathbb{R}^N$  such that*

$$\delta := \text{dist}(A, B) := \inf\{\|x - y\| : x \in A, y \in B\} > 0.$$

*Then  $m_N^*(A \cup B) = m_N^*(A) + m_N^*(B)$ .*

**Theorem 4.1** (4.5). *The Lebesgue measure is a Borel measure, that is,  $\mathcal{B}_N \subseteq \mathcal{M}_N$ .*

## 5 Regularity of the Lebesgue Measure

**Proposition 5.1** (5.1). *For every  $t \in \mathbb{R}^N$  and every  $A \subseteq \mathbb{R}^N$  we have  $m_N^*(t + A) = m_N^*(A)$ .*

**Proposition 5.2** (5.2). *(i) Let  $A \subseteq \mathbb{R}^N$  be an arbitrary set. Then*

$$m_N^*(A) = \inf\{m_N(U) : A \subseteq U, U \text{ open}\}.$$

*(ii) Let  $A \subseteq \mathbb{R}^N$  be Lebesgue measurable. Then*

$$m_N(A) = \sup\{m_N(K) : K \subseteq A, K \text{ compact}\}.$$

**Definition 5.1** (5.3). *A Borel measure on a metric space  $X$  is called outer regular if for every Borel set  $A \subseteq X$*

$$\mu(A) = \inf\{\mu(U) : A \subseteq U, U \text{ open}\}.$$

*The measure is called inner regular if*

$$\mu(A) = \sup\{\mu(K) : K \subseteq A, K \text{ compact}\}.$$

**Corollary 5.1** (5.4). *For every Lebesgue measurable set  $A \subseteq \mathbb{R}^N$  the following assertions are true.*

(i) For every  $\varepsilon > 0$  there exists an open set  $U \subseteq \mathbb{R}^N$  such that  $A \subseteq U$  and

$$m_N(U \setminus A) < \varepsilon.$$

(ii) There exist a sequence  $(U_k)_{k \in \mathbb{N}}$  of open sets with  $A \subseteq U_k$  for all  $k \in \mathbb{N}$  and a set  $S \subseteq A^c$  with  $m_N(S) = 0$  such that

$$B := \bigcap_{k=1}^{\infty} U_k = A \cup S$$

(iii) There exist a sequence  $(C_k)_{k \in \mathbb{N}}$  of compact sets with  $C_k \subseteq A$  for all  $k \in \mathbb{N}$  and a set  $S \subseteq A$  with  $m_N(S) = 0$  such that

$$A = \left( \bigcup_{k=1}^{\infty} C_k \right) \cup S.$$

## 6 Uniqueness of the Lebesgue Measure

**Lemma 6.1** (6.1). Let  $U \subseteq \mathbb{R}^N$  be an open set. Then there exist disjoint dyadic cubes  $Q_j \in \bigcup_{n \in \mathbb{N}} \mathcal{G}_n$  such that  $U = \bigcup_{j=0}^{\infty} Q_j$ .

**Lemma 6.2** (6.2). Let  $\mu$  be a translation invariant Borel measure on  $\mathbb{R}^N$ . If  $Q_{n,k}$  is the dyadic cube given by (6.1), and  $\alpha := \mu((0, 1]^N) < \infty$ , then

$$\mu(Q_{n,k}) = \alpha m_N(Q_{n,k}) = \alpha 2^{-nN}$$

for all  $k \in \mathbb{Z}^N$  and all  $n \in \mathbb{N}$ .

**Theorem 6.1** (6.3). Let  $\mu$  be an outer regular Borel measure on  $\mathbb{R}^N$ . If  $\mu$  is translation invariant and  $\alpha := \mu((0, 1]^N) < \infty$ , then

$$\mu(A) = \alpha m_N(A)$$

for all Borel sets  $A \subseteq \mathbb{R}^N$ .

**Remark 6.1** (6.4). One can show that every Borel measure  $\mu$  on  $\mathbb{R}^N$  with  $\mu(K) < \infty$  for every compact set  $K$  is outer regular.

## 7 Lebesgue Measure and Linear Transformations

**Theorem 7.1** (7.1). Let  $T: \mathbb{R}^N \rightarrow \mathbb{R}^N$  be a linear transformation. Then

$$m_N^*(T(A)) = |\det T| m_N^*(A)$$

for every  $A \subseteq \mathbb{R}^N$ . Moreover, if  $A$  is Lebesgue measurable, then  $T(A)$  is Lebesgue measurable as well.

**Remark 7.1** (7.2). To prove Theorem 7.1 it is in fact sufficient to prove that

$$m_N^*(T(A)) \leq |\det T| m_N^*(A).$$

**Lemma 7.1** (7.3). *Let  $\alpha \neq 0$  and  $A \subseteq \mathbb{R}^N$ . Suppose that*

$$E(x_1, \dots, x_N) := (x_1, \dots, x_{i-1}, \alpha x_i, x_{i+1}, \dots, x_N)$$

*where  $1 \leq i \leq N$ . Then  $m_N^*(E(A)) \leq |\alpha| m_N^*(A) = |\det E| m_N^*(A)$ .*

**Lemma 7.2** (7.4). *Let  $A \subseteq \mathbb{R}^N$  and suppose that*

$$E(x_1, \dots, x_N) := (x_1, \dots, x_{j-1}, x_j + x_i, x_{j+1}, \dots, x_N).$$

*where  $1 \leq i, j \leq N$  and  $i \neq j$ . Then  $m_N^*(E(A)) \leq |\det E| m_N^*(A) = m_N^*(A)$ .*

**Lemma 7.3** (7.5). *Let  $A \subseteq \mathbb{R}^N$  be an arbitrary set and  $T$  an invertible matrix. Then  $m_N^*(T(A)) \leq |\det(T)| m_N^*(A)$ .*

**Lemma 7.4** (7.6). *Suppose that*

$$S_r := \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

*for some  $0 \leq r < N$ . Then  $m_N^*(S_r(A)) = 0$ .*

## 8 The Lebesgue-Stieltjes measure

**Proposition 8.1** (8.1). *Let  $\mu$  be a Borel measure on  $\mathbb{R}$  such that  $\mu([a, b]) < \infty$  for every compact interval  $[a, b] \subseteq \mathbb{R}$ . Then the function  $F: \mathbb{R} \rightarrow [0, \infty)$  given by*

$$F(t) := \begin{cases} \mu((0, t]) & \text{if } t \geq 0, \\ -\mu((t, 0]) & \text{if } t < 0. \end{cases}$$

*is increasing and right continuous, and  $\mu((a, b]) = F(b) - F(a)$  whenever  $a < b$ .*

**Example 8.1** (8.2). (a) *If  $\mu = m_1$  is the Lebesgue measure, then clearly  $F(x) = x$ .*

(b) *Let  $\mu := \delta_0$  be the Dirac measure on  $\mathbb{R}$  concentrated at  $x = 0$ . Then*

$$F(t) = \begin{cases} -1 & \text{if } t < 0, \\ 0 & \text{if } t \geq 0. \end{cases}$$

**Theorem 8.1** (8.3: Lebesgue-Stieltjes measure). *There is a one-to-one correspondence (up to a constant) between right continuous increasing functions and regular Borel measures on  $\mathbb{R}$ :*

(i) *For every regular Borel measure  $\mu$  on  $\mathbb{R}$  there exists a right continuous increasing function  $F: \mathbb{R} \rightarrow \mathbb{R}$  such that  $\mu((a, b]) = F(b) - F(a)$  holds. Moreover, that function is unique up to an additive constant.*

(ii) *For every right continuous increasing function  $F: \mathbb{R} \rightarrow \mathbb{R}$  there exists a unique regular Borel measure such that  $\mu((a, b]) = F(b) - F(a)$  holds.*

**Lemma 8.1** (8.4). *For every  $A \subseteq \mathbb{R}$  we have*

$$\mu_F^*(A) = \inf \left\{ \sum_{k=0}^{\infty} \nu_F((a_k, b_k]) : A \subseteq \bigcup_{k=0}^{\infty} (a_k, b_k] \right\}.$$

**Lemma 8.2** (8.5). *Let  $\mu_F^*$  be the outer measure defined by (8.4). Then*

$$\mu_F^*((a, b]) = F(b) - F(a).$$

**Proposition 8.2** (8.6). *The measure  $\mu_F$  is outer regular.*

## 9 Measurable functions

**Definition 9.1** (9.1). Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $Y$  a topological space (often a subset of  $\mathbb{R}$  or  $\mathbb{C}$ ). We call a function  $f: X \rightarrow Y$   $\mu$ -measurable if  $f^{-1}[U] \in \mathcal{A}$  for all open sets  $U \subseteq Y$ .

**Definition 9.2** (9.4). Let  $X$  be a set and  $A \subseteq X$  a subset. The function  $1_A: X \rightarrow \mathbb{R}$  given by

$$1_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \in X \setminus A \end{cases}$$

is called the indicator function of  $A$ .

**Proposition 9.1** (9.5). Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $A \subseteq X$ . Then  $1_A$  is a measurable function if and only if  $A$  is a measurable set.

**Theorem 9.1** (9.6). Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $f: X \rightarrow [-\infty, \infty]$  a function. Then the following assertions are equivalent:

- (i)  $f$  is measurable;
- (ii)  $f^{-1}[(\alpha, \infty]]$  is measurable for all  $\alpha \in \mathbb{Q}$ ;
- (iii)  $f^{-1}[[\alpha, \infty]]$  is measurable for all  $\alpha \in \mathbb{Q}$ ;
- (iv)  $f^{-1}[[-\infty, \alpha))$  is measurable for all  $\alpha \in \mathbb{Q}$ ;
- (v)  $f^{-1}[[-\infty, \alpha]]$  is measurable for all  $\alpha \in \mathbb{Q}$ .

We can also replace  $\mathbb{Q}$  by some other dense subset of  $\mathbb{R}$  in all the above statements.

**Proposition 9.2** (9.7). Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $f = (f_1, \dots, f_N): X \rightarrow \mathbb{K}^N$  a function. Then  $f$  is measurable if and only if every component function  $f_k: X \rightarrow \mathbb{K}$  is measurable.

**Proposition 9.3** (9.8). Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $Y, Z$  metric spaces. If  $f: X \rightarrow Y$  is measurable and  $\varphi: Y \rightarrow Z$  continuous, then  $\varphi \circ f: X \rightarrow Z$  is measurable.

**Theorem 9.2** (9.9). Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $f, g: X \rightarrow [-\infty, \infty]$  measurable. Then the following functions are measurable as well:

- (i) Let

$$S := X \setminus \{x \in X: f(x) = \infty \text{ and } g(x) = -\infty, \text{ or } f(x) = -\infty \text{ and } g(x) = \infty\}.$$

Then  $(f + g)1_S$  is measurable.

- (ii)  $fg$ ;
- (iii)  $\frac{f}{g}$  if  $g(x) \neq 0$  for all  $x \in X$ ;
- (iv)  $|f|$ ,  $\max\{f, g\}$  and  $\min\{f, g\}$ .

## 10 Sequences of measurable functions

**Proposition 10.1** (10.1). *Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $f_n: X \rightarrow [-\infty, \infty]$  be measurable for all  $n \in \mathbb{N}$ . For  $x \in X$  define*

$$g(x) := \inf_{n \in \mathbb{N}} f_n(x) = \inf\{f_0(x), f_1(x), f_2(x), \dots\}$$

and

$$h(x) := \sup_{n \in \mathbb{N}} f_n(x) = \sup\{f_0(x), f_1(x), f_2(x), \dots\}.$$

Then  $g$  and  $h$  are measurable.

**Theorem 10.1** (10.2). *Let  $(X, \mathcal{A}, \mu)$  be a measure space and let  $f_n: X \rightarrow [-\infty, \infty]$  be measurable functions. Then, the functions  $u$  and  $v$  given by*

$$u(x) = \liminf_{n \rightarrow \infty} f_n(x) \quad \text{and} \quad v(x) = \limsup_{n \rightarrow \infty} f_n(x)$$

are measurable functions. Moreover, if  $f_n \rightarrow f$  pointwise, then  $f$  is measurable.

## 11 Simple measurable functions

**Definition 11.1** (11.1). *Let  $X, Y$  be sets and  $f: X \rightarrow Y$  be a function. We call  $f$  a simple function if its range is a finite set.*

**Proposition 11.1** (11.3). *Suppose that  $f, g: X \rightarrow \mathbb{C}$  are simple measurable functions. Then the following functions are also simple measurable functions:*

- (i)  $f + g$ ;
- (ii)  $\alpha f$  for all  $\alpha \in \mathbb{R}$  (or  $\mathbb{C}$ );
- (iii)  $fg$ ;
- (iv)  $\frac{f}{g}$  if  $g(x) \neq 0$  for all  $x \in X$ .

## 12 Approximation by simple functions

**Theorem 12.1** (12.1). *Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $f: X \rightarrow [0, \infty]$  a measurable function. Then there exists a sequence of simple measurable functions  $\varphi_n: X \rightarrow [0, \infty)$  such that*

$$0 \leq \varphi_n(x) \leq \varphi_{n+1}(x) \leq f(x)$$

for all  $n \in \mathbb{N}$  and all  $x \in X$ . Moreover,  $\varphi_n(x) \rightarrow f(x)$  as  $n \rightarrow \infty$  for all  $x \in X$ , that is,  $\varphi_n \rightarrow f$  pointwise.

**Remark 12.1** (12.2). *If  $f: X \rightarrow [0, \infty)$  is bounded, then  $\varphi_n$  as constructed in the above proof converges uniformly and not just pointwise to  $f$ .*

**Corollary 12.1** (12.3). *Let  $(X, \mathcal{A}, \mu)$  a measure space and  $f: X \rightarrow \mathbb{K}$  measurable. Then there exist simple measurable functions  $\varphi_n: X \rightarrow \mathbb{K}$  such that*

$$0 \leq |\varphi_n(x)| \leq |\varphi_{n+1}(x)| \leq |f(x)|$$

for all  $n \in \mathbb{N}$  and all  $x \in X$ . Moreover,  $\varphi_n \rightarrow f$  pointwise.

## 13 The integration of non-negative simple functions

**Definition 13.1** (14.1). Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $\varphi = \sum_{k=0}^n \alpha_k 1_{A_k}$  be a simple measurable function. We let

$$\int_X \varphi d\mu := \sum_{k=0}^n \alpha_k \mu(A_k).$$

If  $\mu(A_k) = \infty$  we set  $\alpha_k \mu(A_k) = \infty$  if  $\alpha_k > 0$  and  $\alpha_k \mu(A_k) = 0$  if  $\alpha_k = 0$ .

**Proposition 13.1** (14.2). Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $\varphi, \psi$  be simple measurable functions. Then the following assertions are valid.

- (i)  $\int_X \varphi + \psi d\mu = \int_X \varphi d\mu + \int_X \psi d\mu$ ;
- (ii) If  $0 \leq \varphi \leq \psi$ , then  $\int_X \varphi d\mu \leq \int_X \psi d\mu$ ;
- (iii)  $\alpha \int_X \varphi d\mu = \int_X \alpha \varphi d\mu$  for all  $\alpha \geq 0$ ;
- (iv) If  $N \in \mathcal{A}$  and  $\mu(N) = 0$ , then  $\int_X \varphi d\mu = \int_X 1_{X \setminus N} \varphi d\mu$ .

## 14 Integration of non-negative measurable functions

**Definition 14.1** (15.1). Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $f: X \rightarrow [0, \infty]$  a measurable function. We then set

$$\int_X f d\mu = \sup \left\{ \int_X \varphi d\mu : \varphi \text{ is simple measurable and } 0 \leq \varphi \leq f \right\}.$$

**Definition 14.2** (15.4). Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $(P)$  some property. We say that  $(P)$  holds almost everywhere on  $X$  if there exists a measurable set  $N$  such that  $\mu(N) = 0$  and  $(P)$  holds for all  $x \in X \setminus N$ .

**Lemma 14.1** (15.5: Markov's inequality). Suppose that  $f: X \rightarrow [0, \infty]$  is measurable. Then for every  $\alpha > 0$  we have that

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

**Proposition 14.1** (15.6). Suppose that  $f: X \rightarrow [0, \infty]$  is measurable.

- (i) If  $\int_X f d\mu < \infty$ , then  $f(x) < \infty$  almost everywhere, that is,

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

- (ii) If  $\int_X f d\mu = 0$ , then  $f(x) = 0$  almost everywhere, that is,

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

**Remark 14.1** (15.7). From the above proposition, if the integral of a non-negative function  $f$  is zero, then we cannot conclude in general that  $f$  is zero.

## 15 The monotone convergence theorem

**Theorem 15.1** (16.1: Monotone convergence theorem). *Let  $(X, \mathcal{A}, \mu)$  be a measure space. For every  $n \in \mathbb{N}$  let*

$$f_n: X \rightarrow [0, \infty]$$

*be a measurable function and suppose that*

$$0 \leq f_n(x) \leq f_{n+1}(x)$$

*for almost every  $x \in X$ . Then there exists a measurable function  $f: X \rightarrow [0, \infty]$  such that  $f_n(x) \rightarrow f(x)$  as  $n \rightarrow \infty$  for almost every  $x \in X$  and*

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

**Theorem 15.2** (16.2: Fatou's Lemma). *Let  $(X, \mathcal{A}, \mu)$  be a measure space. For every  $n \in \mathbb{N}$  let*

$$f_n: X \rightarrow [0, \infty]$$

*be measurable functions. Then*

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

**Proposition 15.1** (16.3). *Suppose that  $f, g: X \rightarrow [0, \infty]$  are measurable functions and that  $\alpha \geq 0$  is a constant. Then*

$$(i) \int_X f + g d\mu = \int_X f d\mu + \int_X g d\mu;$$

$$(ii) \alpha \int_X f d\mu = \int_X \alpha f d\mu.$$

**Theorem 15.3** (16.4). *For every  $k \in \mathbb{N}$  let  $g_k: X \rightarrow [0, \infty]$  be measurable. Then*

$$\sum_{k=0}^{\infty} \int_X g_k d\mu = \int_X \sum_{k=0}^{\infty} g_k d\mu.$$

## 16 Integrable functions

**Definition 16.1** (17.1). *Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $f: X \rightarrow \mathbb{K}$  measurable (with  $\mathbb{K} = \mathbb{R}$  or  $\mathbb{C}$  as usual). We call  $f$  a  $\mu$ -integrable function if*

$$\int_X |f| d\mu < \infty.$$

*We let*

$$\mathcal{L}^1(X, \mathcal{A}, \mu; \mathbb{K}) := \{f: X \rightarrow \mathbb{K} \mid f \text{ is } \mu\text{-integrable}\}$$

*and call it the space of integrable functions.*

**Definition 16.2** (17.3). (i) For  $f \in \mathcal{L}^1(X, \mathbb{R})$  we define

$$\int_X f d\mu := \int_X f^+ d\mu - \int_X f^- d\mu.$$

(ii) For  $f \in \mathcal{L}^1(X, \mathbb{C})$  we define

$$\int_X f d\mu := \int_X \operatorname{Re} f d\mu + i \int_X \operatorname{Im} f d\mu.$$

(iii) For any  $\mu$ -measurable set  $A$  and  $f \in \mathcal{L}^1(X, \mathbb{K})$  we define

$$\int_A f d\mu := \int_X 1_A f d\mu.$$

**Theorem 16.1** (17.4). Let  $f, g \in \mathcal{L}^1(X, \mathcal{A}, \mu; \mathbb{K})$ . Then the following assertions are true.

(i) for all  $\alpha, \beta \in \mathbb{K}$

$$\int_X \alpha f + \beta g d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu.$$

(ii) If  $A, B \in \mathcal{A}$  are disjoint, then

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

(iii) We have the "triangle inequality"

$$\left| \int_X f d\mu \right| \leq \int_X |f| d\mu.$$

## 17 Limit Theorems

**Theorem 17.1** (20.1: Dominated convergence theorem). Let  $f_n: X \rightarrow \mathbb{K}$  be measurable and  $f: X \rightarrow \mathbb{K}$  be such that  $f_n(x) \rightarrow f(x)$  pointwise almost everywhere. Furthermore assume that there exists  $g \in \mathcal{L}^1(X)$  such that

$$|f_n(x)| \leq g(x)$$

for all  $n \in \mathbb{N}$  and for almost all  $x \in X$ . Then

(i)  $f_n, f \in \mathcal{L}^1(X, \mathbb{K})$  for all  $n \in \mathbb{N}$ ;

(ii)  $\int_X |f_n - f| d\mu \rightarrow 0$  as  $n \rightarrow \infty$ ;

(iii)  $\int_X f_n d\mu \rightarrow \int_X f d\mu$  as  $n \rightarrow \infty$ .

**Corollary 17.1** (20.2). Let  $f \in \mathcal{L}^1((a, b), \mathbb{K})$ . Then

$$\lim_{x \rightarrow b^-} \int_a^x f(t) dt = \int_a^b f(t) dt$$

**Theorem 17.2** (20.3: Continuity of parameter integrals). *Let  $(X, \mathcal{A}, \mu)$  a measure space and  $Y$  a metric space (usually a subset of  $\mathbb{R}$  or  $\mathbb{C}$ ). Suppose that  $f: X \times Y \rightarrow \mathbb{K}$  such that*

- $x \mapsto f(x, y)$  is  $\mu$ -measurable for all  $y \in Y$ ;
- $y \mapsto f(x, y)$  is continuous at  $y_0$  for almost all  $x \in X$ ;
- there exists  $g \in L^1(X, \mathbb{R})$  such that

$$|f(x, y)| \leq g(x)$$

for almost all  $x \in X$  and all  $y \in Y$ .

For  $y \in Y$  define

$$F(y) := \int_X f(x, y) d\mu.$$

Then  $F$  is continuous at  $y_0$ .

**Theorem 17.3** (20.4: Differentiation of parameter integrals). *Let  $(X, \mathcal{A}, \mu)$  a measure space and  $J \subseteq \mathbb{R}$  an interval. Suppose that  $f: X \times J \rightarrow \mathbb{K}$  is a function such that*

- $x \mapsto f(x, t)$  is  $\mu$ -integrable for all  $t \in J$ ;
- for almost all  $x \in X$ ,  $\frac{\partial f}{\partial t}(x, t)$  exists and is continuous on  $J$ ;
- there exists  $g \in \mathcal{L}^1(X, \mathbb{R})$  such that

$$\left| \frac{\partial f}{\partial t}(x, t) \right| \leq g(x)$$

for almost all  $x \in X$  and all  $t \in J$ .

For  $t \in J$  define

$$F(t) := \int_X f(x, t) d\mu.$$

Then  $F: J \rightarrow \mathbb{K}$  is differentiable and

$$F'(t) = \int_X \frac{\partial f}{\partial t}(x, t) d\mu$$

for all  $t \in J$ .

**Theorem 17.4** (20.5: Fundamental Theorem of Calculus). (i) *If  $f: [a, b] \rightarrow \mathbb{R}$  is continuous, then*

$$\frac{d}{dt} \int_a^t f(s) ds = f(t)$$

for all  $t \in [a, b]$ .

(ii) *If  $f: [a, b] \rightarrow \mathbb{R}$  is continuous and  $f: (a, b) \rightarrow \mathbb{R}$  continuously differentiable, then*

$$\int_a^b f'(s) ds = f(b) - f(a).$$

## 18 The $\mathcal{L}^p$ -spaces

**Definition 18.1** (21.1). Let  $1 \leq p < \infty$  and  $f: X \rightarrow \mathbb{K}$ . We call

$$\|f\|_p := \left( \int_X |f|^p d\mu \right)^{1/p}$$

the  $L^p$ -norm of  $f$ . We set

$$\mathcal{L}^p(X, \mathcal{A}, \mu, \mathbb{K}) := \{f: X \rightarrow \mathbb{K} \mid f \text{ measurable and } \|f\|_p < \infty\}$$

**Lemma 18.1** (21.3: Young's inequality). Let  $p, q > 1$  such that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then

$$st \leq \frac{1}{p}s^p + \frac{1}{q}t^q$$

for all  $s, t \geq 0$ .

**Theorem 18.1** (21.4: Hölder's inequality). Let  $p, q > 1$  such that  $\frac{1}{p} + \frac{1}{q} = 1$ . Then

$$\int_X |fg| d\mu \leq \|f\|_p \|g\|_q$$

for all  $f \in \mathcal{L}^p(X)$  and  $g \in \mathcal{L}^q(X)$ .

**Proposition 18.1** (21.5). Let  $1 \leq p < q < \infty$ . If  $\mu(X) < \infty$ , then  $\mathcal{L}^q(X) \subseteq \mathcal{L}^p(X)$ . More precisely, if  $f \in \mathcal{L}^q(X)$ , then

$$\|f\|_p \leq \mu(X)^{\frac{1}{p} - \frac{1}{q}} \|f\|_q.$$

In general,  $\mathcal{L}^p(X) \not\subseteq \mathcal{L}^q(X)$ . Moreover, if  $\mu(X) = \infty$ , then neither  $\mathcal{L}^p(X) \subseteq \mathcal{L}^q(X)$  nor  $\mathcal{L}^q(X) \subseteq \mathcal{L}^p(X)$  in general.

**Proposition 18.2** (21.6: Minkowski's inequality). For  $1 \leq p < \infty$  we have

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

for all  $f, g \in \mathcal{L}^p(X)$ .

**Theorem 18.2** (21.7: Properties of  $L^p$ -norm). Let  $1 \leq p < \infty$ . Then for  $f, g \in \mathcal{L}^p(X)$  and  $\alpha \in \mathbb{K}$  we have

- (i)  $\|f\|_p \geq 0$  with equality if and only if  $f = 0$  almost everywhere.
- (ii)  $\|\alpha f\|_p = |\alpha| \|f\|_p$
- (iii)  $\|f + g\|_p \leq \|f\|_p + \|g\|_p$  (triangle or Minkowski's inequality)

**Definition 18.2** (21.8). Let  $f_n, f \in \mathcal{L}^p(X)$ .

- (i) We say that  $f_n \rightarrow f$  in  $\mathcal{L}^p(X)$  if  $\|f_n - f\|_p \rightarrow 0$  as  $n \rightarrow \infty$ .

(ii) We say that  $(f_n)$  is a Cauchy sequence in  $\mathcal{L}^p(X)$  if for every  $\varepsilon > 0$  there exists  $n_0 \in \mathbb{N}$  such that

$$\|f_n - f_m\|_p < \varepsilon$$

for all  $n, m > n_0$ .

**Lemma 18.2** (21.10). Suppose that  $(f_n)$  is a Cauchy sequence in  $\mathcal{L}^p(X)$ . If  $(f_{n_k})$  is a convergent subsequence with  $f_{n_k} \rightarrow f$  in  $\mathcal{L}^p(X)$ , then  $f_n \rightarrow f$  in  $\mathcal{L}^p(X)$ .

**Proposition 18.3** (21.11). Let  $(g_k)$  be a sequence in  $\mathcal{L}^p(X)$  such that

$$\sum_{k=1}^{\infty} \|g_k\|_p < \infty.$$

Then there exists a function  $f \in \mathcal{L}^p(X)$  such that  $f = \sum_{k=0}^{\infty} g_k$  converges pointwise almost everywhere and in  $\mathcal{L}^p(X)$ .

**Theorem 18.3** (21.12: Completeness of  $\mathcal{L}^p$ ). Let  $(f_n)$  be a Cauchy sequence in  $\mathcal{L}^p(X)$ . Then there exists a function  $f \in \mathcal{L}^p(X)$  such that  $f_n \rightarrow f$  in  $\mathcal{L}^p(X)$ . Moreover,  $(f_n)$  has a subsequence that converges pointwise almost everywhere.

**Remark 18.1** (21.13). (a) In the above theorem we did not claim that  $f_n \rightarrow f$  pointwise almost everywhere.

(b) If we know that  $f_n \rightarrow g$  pointwise almost everywhere and  $f_n \rightarrow f$  in  $\mathcal{L}^p(X)$ , then we can conclude that  $f = g$  almost everywhere.

(c) If we happen to know that  $f_n \rightarrow f$  in  $\mathcal{L}^p(X)$  and the increments  $\|f_k - f_{k-1}\|_p$  converge "fast enough", then  $f_n \rightarrow f$  pointwise almost everywhere.

## 19 The $L^p$ -spaces

**Definition 19.1** (22.1). For  $1 \leq p < \infty$  we define

$$L^p(X, \mathcal{A}, \mu, \mathbb{K}) := \{f : f \in \mathcal{L}^p(X, \mathcal{A}, \mu, \mathbb{K})\}$$

We set

$$\|f\|_p := \|f\|_p.$$

Similarly we set

$$f + g := f + g \quad \text{and} \quad \alpha f := \alpha f$$

for all  $f, g \in L^p(X)$  and  $\alpha \in \mathbb{K}$ .

**Theorem 19.1** (22.4). For  $1 \leq p < \infty$ , the space  $L^p(X)$  is a complete normed space with respect to the norm  $\|\cdot\|_p$ .

**Remark 19.1** (22.5). (a) A complete normed vector space is called a Banach space, and the Lebesgue spaces  $L^p(X)$  is one of the most important example of such spaces.

(b) The case  $p = 2$  plays an important role since then the  $L^2$ -norm is induced by the inner product

$$(f | g) = \int_X f \bar{g} d\mu$$

## 20 Basic density theorems

**Theorem 20.1** (23.1: Density of simple functions in  $L^p$ ). *Let  $(X, \mathcal{A}, \mu)$  be a measure space and  $1 \leq p < \infty$ . Then the simple functions are dense in  $L^p(X)$ . That is, for every  $f \in L^p(X)$  and every  $\varepsilon > 0$ , there exists a simple function  $\varphi$  such that*

$$\|f - \varphi\|_p < \varepsilon.$$

**Theorem 20.2** (23.2: Density of continuous functions in  $L^p(\mathbb{R}^N)$ ). *Let  $1 \leq p < \infty$ . Then the space  $C_c(\mathbb{R}^N)$  of continuous functions with compact support is dense in  $L^p(\mathbb{R}^N)$ .*

**Theorem 20.3** (23.3: Density of step functions in  $L^p(\mathbb{R}^N)$ ). *Let  $1 \leq p < \infty$ . Then the step functions (finite linear combinations of characteristic functions of rectangles) are dense in  $L^p(\mathbb{R}^N)$ .*

## 21 The space of bounded measurable functions

**Definition 21.1** (24.1). *Let  $(X, \mathcal{A}, \mu)$  be a measure space. We denote by  $\mathcal{B}(X)$  the space of all bounded measurable functions  $f : X \rightarrow \mathbb{K}$ , equipped with the supremum norm*

$$\|f\|_\infty := \sup\{|f(x)| : x \in X\}.$$

**Proposition 21.1** (24.2). *The space  $\mathcal{B}(X)$  is a Banach space with respect to the supremum norm  $\|\cdot\|_\infty$ .*

**Definition 21.2** (24.3). *For  $f \in \mathcal{L}^\infty(X)$ , we define the essential supremum norm by*

$$\|f\|_\infty := \inf\{M \geq 0 : |f(x)| \leq M \text{ for almost every } x \in X\}.$$

*The space  $L^\infty(X)$  consists of equivalence classes of essentially bounded measurable functions.*

**Theorem 21.1** (24.4: Properties of  $L^\infty$ ). *The space  $L^\infty(X)$  is a Banach space with respect to the essential supremum norm  $\|\cdot\|_\infty$ .*

**Proposition 21.2** (24.5: Hölder's inequality for  $p = \infty$ ). *If  $f \in L^1(X)$  and  $g \in L^\infty(X)$ , then  $fg \in L^1(X)$  and*

$$\|fg\|_1 \leq \|f\|_1 \|g\|_\infty.$$

## 22 Fubini's Theorem

**Definition 22.1** (25.1). *Let  $(X, \mathcal{A}, \mu)$  and  $(Y, \mathcal{B}, \nu)$  be measure spaces. The product  $\sigma$ -algebra  $\mathcal{A} \otimes \mathcal{B}$  is the  $\sigma$ -algebra generated by all sets of the form  $A \times B$  with  $A \in \mathcal{A}$  and  $B \in \mathcal{B}$ .*

**Theorem 22.1** (25.2: Existence of product measure). *There exists a unique measure  $\mu \otimes \nu$  on  $(X \times Y, \mathcal{A} \otimes \mathcal{B})$  such that*

$$(\mu \otimes \nu)(A \times B) = \mu(A)\nu(B)$$

*for all  $A \in \mathcal{A}$ ,  $B \in \mathcal{B}$ .*

**Theorem 22.2** (25.3: Fubini's Theorem). *Let  $(X, \mathcal{A}, \mu)$  and  $(Y, \mathcal{B}, \nu)$  be  $\sigma$ -finite measure spaces, and let  $f : X \times Y \rightarrow \mathbb{K}$  be  $\mu \otimes \nu$ -integrable. Then:*

- (i) *For almost every  $x \in X$ , the function  $y \mapsto f(x, y)$  is  $\nu$ -integrable.*
- (ii) *For almost every  $y \in Y$ , the function  $x \mapsto f(x, y)$  is  $\mu$ -integrable.*
- (iii) *The functions*

$$x \mapsto \int_Y f(x, y) d\nu(y) \quad \text{and} \quad y \mapsto \int_X f(x, y) d\mu(x)$$

*are integrable on  $X$  and  $Y$  respectively.*

(iv)

$$\int_{X \times Y} f d(\mu \otimes \nu) = \int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left( \int_X f(x, y) d\mu(x) \right) d\nu(y).$$

**Theorem 22.3** (25.4: Tonelli's Theorem). *Let  $(X, \mathcal{A}, \mu)$  and  $(Y, \mathcal{B}, \nu)$  be  $\sigma$ -finite measure spaces, and let  $f : X \times Y \rightarrow [0, \infty]$  be measurable. Then:*

- (i) *For almost every  $x \in X$ , the function  $y \mapsto f(x, y)$  is measurable.*
- (ii) *For almost every  $y \in Y$ , the function  $x \mapsto f(x, y)$  is measurable.*
- (iii) *The functions*

$$x \mapsto \int_Y f(x, y) d\nu(y) \quad \text{and} \quad y \mapsto \int_X f(x, y) d\mu(x)$$

*are measurable.*

(iv)

$$\int_{X \times Y} f d(\mu \otimes \nu) = \int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left( \int_X f(x, y) d\mu(x) \right) d\nu(y).$$

## 23 Translation of a function

**Definition 23.1** (26.1). *For  $h \in \mathbb{R}^N$  and a function  $f : \mathbb{R}^N \rightarrow \mathbb{K}$ , we define the translation of  $f$  by  $h$  as*

$$(\tau_h f)(x) := f(x - h).$$

**Proposition 23.1** (26.2: Properties of translation). *Let  $f, g \in L^p(\mathbb{R}^N)$ ,  $1 \leq p \leq \infty$ , and  $h, k \in \mathbb{R}^N$ . Then:*

- (i)  $\tau_h(\tau_k f) = \tau_{h+k} f$
- (ii)  $\|\tau_h f\|_p = \|f\|_p$
- (iii)  $\lim_{h \rightarrow 0} \|\tau_h f - f\|_p = 0$  for  $1 \leq p < \infty$

**Definition 23.2** (26.3). *A linear operator  $T : L^p(\mathbb{R}^N) \rightarrow L^p(\mathbb{R}^N)$  is called translation invariant if*

$$T(\tau_h f) = \tau_h(Tf)$$

*for all  $f \in L^p(\mathbb{R}^N)$  and  $h \in \mathbb{R}^N$ .*

## 24 Convex functions and Jensen's inequality

**Definition 24.1** (27.1). A function  $\varphi : (a, b) \rightarrow \mathbb{R}$  is called convex if for all  $x, y \in (a, b)$  and  $\lambda \in [0, 1]$ ,

$$\varphi(\lambda x + (1 - \lambda)y) \leq \lambda\varphi(x) + (1 - \lambda)\varphi(y).$$

**Proposition 24.1** (27.2). If  $\varphi$  is twice differentiable on  $(a, b)$  and  $\varphi''(x) \geq 0$  for all  $x \in (a, b)$ , then  $\varphi$  is convex.

**Theorem 24.1** (27.3: Jensen's inequality). Let  $(X, \mathcal{A}, \mu)$  be a probability space (i.e.,  $\mu(X) = 1$ ), and let  $f \in L^1(X, \mathbb{R})$  be such that  $f(x) \in (a, b)$  for almost every  $x \in X$ . If  $\varphi : (a, b) \rightarrow \mathbb{R}$  is convex, then

$$\varphi\left(\int_X f d\mu\right) \leq \int_X (\varphi \circ f) d\mu.$$

**Corollary 24.1** (27.4). For  $f \in L^p(X)$ ,  $1 \leq p < \infty$ , we have

$$\left|\int_X f d\mu\right|^p \leq \int_X |f|^p d\mu.$$

## 25 Convolution

**Definition 25.1** (28.1). For  $f, g \in L^1(\mathbb{R}^N)$ , the convolution of  $f$  and  $g$  is defined by

$$(f * g)(x) := \int_{\mathbb{R}^N} f(x - y)g(y) dy.$$

**Theorem 25.1** (28.2: Properties of convolution). Let  $f, g, h \in L^1(\mathbb{R}^N)$ . Then:

- (i)  $f * g \in L^1(\mathbb{R}^N)$  and  $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$
- (ii)  $f * g = g * f$  (commutativity)
- (iii)  $(f * g) * h = f * (g * h)$  (associativity)
- (iv)  $f * (g + h) = f * g + f * h$  (distributivity)

**Theorem 25.2** (28.3: Young's inequality for convolution). Let  $1 \leq p, q, r \leq \infty$  satisfy

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}.$$

If  $f \in L^p(\mathbb{R}^N)$  and  $g \in L^q(\mathbb{R}^N)$ , then  $f * g \in L^r(\mathbb{R}^N)$  and

$$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

**Proposition 25.1** (28.4: Smoothing property of convolution). If  $f \in L^1(\mathbb{R}^N)$  and  $g \in C^k(\mathbb{R}^N)$  with bounded derivatives, then  $f * g \in C^k(\mathbb{R}^N)$  and

$$\partial^\alpha (f * g) = f * (\partial^\alpha g)$$

for all multi-indices  $\alpha$  with  $|\alpha| \leq k$ .

## 26 Approximate identities

**Definition 26.1** (29.1). A family  $\{\varphi_\varepsilon\}_{\varepsilon>0}$  of functions in  $L^1(\mathbb{R}^N)$  is called an approximate identity if:

- (i)  $\int_{\mathbb{R}^N} \varphi_\varepsilon(x) dx = 1$  for all  $\varepsilon > 0$
- (ii)  $\sup_{\varepsilon>0} \|\varphi_\varepsilon\|_1 < \infty$
- (iii) For every  $\delta > 0$ ,  $\lim_{\varepsilon \rightarrow 0} \int_{|x|>\delta} |\varphi_\varepsilon(x)| dx = 0$

**Theorem 26.1** (29.2: Approximation by convolution). Let  $\{\varphi_\varepsilon\}$  be an approximate identity and  $f \in L^p(\mathbb{R}^N)$  for  $1 \leq p < \infty$ . Then

$$\lim_{\varepsilon \rightarrow 0} \|f * \varphi_\varepsilon - f\|_p = 0.$$

If  $f$  is bounded and uniformly continuous, then  $f * \varphi_\varepsilon \rightarrow f$  uniformly as  $\varepsilon \rightarrow 0$ .

**Example 26.1** (29.3). The Poisson kernel on  $\mathbb{R}$ :

$$P_\varepsilon(x) = \frac{1}{\pi} \frac{\varepsilon}{x^2 + \varepsilon^2}$$

is an approximate identity.

**Example 26.2** (29.4). The Gauss-Weierstrass kernel:

$$G_\varepsilon(x) = (4\pi\varepsilon)^{-N/2} e^{-|x|^2/(4\varepsilon)}$$

is an approximate identity.

## 27 Approximation theorems

**Theorem 27.1** (30.1: Meyers-Serrin Theorem). Let  $\Omega \subseteq \mathbb{R}^N$  be open and  $1 \leq p < \infty$ . Then  $C^\infty(\Omega) \cap L^p(\Omega)$  is dense in  $L^p(\Omega)$ .

**Theorem 27.2** (30.2: Density of smooth functions with compact support). Let  $\Omega \subseteq \mathbb{R}^N$  be open and  $1 \leq p < \infty$ . Then  $C_c^\infty(\Omega)$  is dense in  $L^p(\Omega)$ .

**Theorem 27.3** (30.3: Urysohn's Lemma). Let  $X$  be a normal topological space, and let  $A, B$  be disjoint closed subsets of  $X$ . Then there exists a continuous function  $f : X \rightarrow [0, 1]$  such that  $f|_A = 0$  and  $f|_B = 1$ .

**Theorem 27.4** (30.4: Partition of unity). Let  $X$  be a locally compact Hausdorff space, and let  $\{U_\alpha\}_{\alpha \in I}$  be an open cover of  $X$ . Then there exists a family  $\{\varphi_\alpha\}_{\alpha \in I}$  of continuous functions  $\varphi_\alpha : X \rightarrow [0, 1]$  such that:

- (i) Each  $\varphi_\alpha$  has compact support contained in  $U_\alpha$
- (ii) The family  $\{\text{supp}(\varphi_\alpha)\}$  is locally finite
- (iii)  $\sum_{\alpha \in I} \varphi_\alpha(x) = 1$  for all  $x \in X$

## 28 Definition and basic properties of the Fourier transform

**Definition 28.1** (31.1). For  $f \in L^1(\mathbb{R}^N)$ , the Fourier transform of  $f$  is defined by

$$\hat{f}(\xi) = \mathcal{F}f(\xi) := \int_{\mathbb{R}^N} f(x) e^{-2\pi i x \cdot \xi} dx.$$

**Proposition 28.1** (31.2: Basic properties of Fourier transform). Let  $f, g \in L^1(\mathbb{R}^N)$ .

- (i)  $\|\hat{f}\|_\infty \leq \|f\|_1$
- (ii)  $\hat{f}$  is uniformly continuous on  $\mathbb{R}^N$
- (iii)  $\widehat{f * g} = \hat{f} \cdot \hat{g}$
- (iv)  $\widehat{\tau_h f}(\xi) = e^{-2\pi i h \cdot \xi} \hat{f}(\xi)$
- (v)  $\widehat{e^{2\pi i h \cdot x} f}(x)(\xi) = \hat{f}(\xi - h)$

**Theorem 28.1** (31.3: Riemann-Lebesgue Lemma). If  $f \in L^1(\mathbb{R}^N)$ , then  $\hat{f} \in C_0(\mathbb{R}^N)$ , i.e.,  $\hat{f}$  is continuous and  $\lim_{|\xi| \rightarrow \infty} \hat{f}(\xi) = 0$ .

**Proposition 28.2** (31.4: Fourier transform of derivatives). If  $f \in C^1(\mathbb{R}^N) \cap L^1(\mathbb{R}^N)$  and  $\partial_j f \in L^1(\mathbb{R}^N)$ , then

$$\widehat{\partial_j f}(\xi) = 2\pi i \xi_j \hat{f}(\xi).$$

**Proposition 28.3** (31.5: Derivative of Fourier transform). If  $f \in L^1(\mathbb{R}^N)$  and  $x_j f(x) \in L^1(\mathbb{R}^N)$ , then  $\hat{f}$  is differentiable with respect to  $\xi_j$  and

$$\partial_{\xi_j} \hat{f}(\xi) = -2\pi i \widehat{x_j f}(x)(\xi).$$

## 29 Fundamental properties of the Fourier transform

**Theorem 29.1** (32.1: Fourier inversion theorem). Let  $f \in L^1(\mathbb{R}^N)$  and suppose  $\hat{f} \in L^1(\mathbb{R}^N)$ . Then for almost every  $x \in \mathbb{R}^N$ ,

$$f(x) = \int_{\mathbb{R}^N} \hat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi.$$

If  $f$  is continuous, the equality holds everywhere.

**Definition 29.1** (32.2). The inverse Fourier transform is defined by

$$\check{f}(x) = \mathcal{F}^{-1}f(x) := \int_{\mathbb{R}^N} f(\xi) e^{2\pi i x \cdot \xi} d\xi.$$

**Theorem 29.2** (32.3: Plancherel's theorem). The Fourier transform extends uniquely to a unitary operator on  $L^2(\mathbb{R}^N)$ . That is, for  $f \in L^2(\mathbb{R}^N)$ ,

$$\|\hat{f}\|_2 = \|f\|_2.$$

Moreover, the Fourier transform is a bijection on  $L^2(\mathbb{R}^N)$  with inverse  $\mathcal{F}^{-1}$ .

**Theorem 29.3** (32.4: Parseval's identity). For  $f, g \in L^2(\mathbb{R}^N)$ ,

$$\int_{\mathbb{R}^N} f(x) \overline{g(x)} dx = \int_{\mathbb{R}^N} \hat{f}(\xi) \overline{\hat{g}(\xi)} d\xi.$$

**Proposition 29.1** (32.5: Fourier transform of Gaussian). The Fourier transform of the Gaussian function  $f(x) = e^{-\pi|x|^2}$  is

$$\hat{f}(\xi) = e^{-\pi|\xi|^2}.$$

**Theorem 29.4** (32.6: Convolution theorem). For  $f, g \in L^1(\mathbb{R}^N)$ ,

$$\widehat{f * g} = \hat{f} \cdot \hat{g}.$$

For  $f, g \in L^2(\mathbb{R}^N)$ ,

$$\widehat{f \cdot g} = \hat{f} * \hat{g}.$$

## 30 The Fourier transform on $L^2(\mathbb{R}^N)$

**Definition 30.1** (33.1). For  $f \in L^2(\mathbb{R}^N)$ , the Fourier transform is defined by

$$\hat{f}(\xi) = \lim_{R \rightarrow \infty} \int_{|x| \leq R} f(x) e^{-2\pi i x \cdot \xi} dx,$$

where the limit is in the  $L^2$ -sense.

**Theorem 30.1** (33.2). The Fourier transform  $\mathcal{F} : L^2(\mathbb{R}^N) \rightarrow L^2(\mathbb{R}^N)$  is a unitary operator, i.e., it is bijective and preserves the inner product:

$$\langle f, g \rangle = \langle \hat{f}, \hat{g} \rangle \quad \text{for all } f, g \in L^2(\mathbb{R}^N).$$

**Proposition 30.1** (33.3). The Fourier transform maps the Schwartz space  $\mathcal{S}(\mathbb{R}^N)$  onto itself.

**Definition 30.2** (33.4). The Schwartz space  $\mathcal{S}(\mathbb{R}^N)$  consists of all infinitely differentiable functions  $f : \mathbb{R}^N \rightarrow \mathbb{C}$  such that for all multi-indices  $\alpha, \beta$ ,

$$\sup_{x \in \mathbb{R}^N} |x^\alpha \partial^\beta f(x)| < \infty.$$

**Theorem 30.2** (33.5). The Fourier transform is an isomorphism of the Schwartz space  $\mathcal{S}(\mathbb{R}^N)$  onto itself.

## 31 The Riesz representation theorem

**Definition 31.1** (34.1). Let  $X$  be a topological space. A linear functional  $\Lambda : C_c(X) \rightarrow \mathbb{C}$  is called positive if  $\Lambda(f) \geq 0$  whenever  $f \geq 0$ .

**Theorem 31.1** (34.2: Riesz-Markov-Kakutani representation theorem). Let  $X$  be a locally compact Hausdorff space, and let  $\Lambda : C_c(X) \rightarrow \mathbb{C}$  be a positive linear functional. Then there exists a unique Radon measure  $\mu$  on  $X$  such that

$$\Lambda(f) = \int_X f d\mu \quad \text{for all } f \in C_c(X).$$

**Definition 31.2** (34.3). A Radon measure on a locally compact Hausdorff space  $X$  is a Borel measure that is:

- (i) Finite on compact sets
- (ii) Outer regular on all Borel sets
- (iii) Inner regular on open sets

**Theorem 31.2** (34.4: Riesz representation theorem for  $C_0(X)$ ). Let  $X$  be a locally compact Hausdorff space, and let  $\Lambda : C_0(X) \rightarrow \mathbb{C}$  be a continuous linear functional. Then there exists a unique complex Radon measure  $\mu$  on  $X$  such that

$$\Lambda(f) = \int_X f d\mu \quad \text{for all } f \in C_0(X),$$

and  $\|\Lambda\| = |\mu|(X)$ .

## 32 The Radon-Nikodym theorem

**Definition 32.1** (35.1). Let  $\mu$  and  $\nu$  be measures on  $(X, \mathcal{A})$ . We say that  $\nu$  is absolutely continuous with respect to  $\mu$  (written  $\nu \ll \mu$ ) if  $\mu(A) = 0$  implies  $\nu(A) = 0$  for all  $A \in \mathcal{A}$ .

**Definition 32.2** (35.2). A measure  $\nu$  is  $\sigma$ -finite if  $X$  can be written as a countable union of sets with finite  $\nu$ -measure.

**Theorem 32.1** (35.3: Radon-Nikodym theorem). Let  $\mu$  and  $\nu$  be  $\sigma$ -finite measures on  $(X, \mathcal{A})$  with  $\nu \ll \mu$ . Then there exists a measurable function  $f : X \rightarrow [0, \infty)$  such that

$$\nu(A) = \int_A f d\mu \quad \text{for all } A \in \mathcal{A}.$$

The function  $f$  is unique up to equality  $\mu$ -almost everywhere, and is called the Radon-Nikodym derivative, denoted  $f = \frac{d\nu}{d\mu}$ .

**Theorem 32.2** (35.4: Lebesgue decomposition theorem). Let  $\mu$  and  $\nu$  be  $\sigma$ -finite measures on  $(X, \mathcal{A})$ . Then there exist unique measures  $\nu_a$  and  $\nu_s$  such that:

- (i)  $\nu = \nu_a + \nu_s$
- (ii)  $\nu_a \ll \mu$
- (iii)  $\nu_s \perp \mu$  (i.e., there exists  $E \in \mathcal{A}$  with  $\mu(E) = 0$  and  $\nu_s(X \setminus E) = 0$ )

**Proposition 32.1** (35.5: Chain rule for Radon-Nikodym derivatives). If  $\lambda \ll \nu \ll \mu$  and  $\nu$  and  $\mu$  are  $\sigma$ -finite, then

$$\frac{d\lambda}{d\mu} = \frac{d\lambda}{d\nu} \frac{d\nu}{d\mu} \quad \mu\text{-a.e.}$$

### 33 Basic concepts of probability theory

**Definition 33.1** (36.1). A probability space is a measure space  $(\Omega, \mathcal{F}, P)$  with  $P(\Omega) = 1$ . The set  $\Omega$  is called the sample space,  $\mathcal{F}$  is called the  $\sigma$ -algebra of events, and  $P$  is called the probability measure.

**Definition 33.2** (36.2). A random variable is a measurable function  $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ .

**Definition 33.3** (36.3). The distribution of a random variable  $X$  is the probability measure  $\mu_X$  on  $\mathbb{R}$  defined by

$$\mu_X(B) = P(X \in B) = P(\{\omega \in \Omega : X(\omega) \in B\}).$$

**Definition 33.4** (36.4). The expectation of a random variable  $X$  is defined by

$$\mathbb{E}[X] = \int_{\Omega} X dP,$$

provided the integral exists.

**Definition 33.5** (36.5). The variance of a random variable  $X$  is defined by

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

**Theorem 33.1** (36.6: Law of the unconscious statistician). If  $X$  is a random variable and  $g : \mathbb{R} \rightarrow \mathbb{R}$  is measurable, then

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x) d\mu_X(x).$$

### 34 The distribution of a random variable

**Definition 34.1** (37.1). The cumulative distribution function (CDF) of a random variable  $X$  is defined by

$$F_X(x) = P(X \leq x) = \mu_X((-\infty, x]).$$

**Proposition 34.1** (37.2). The cumulative distribution function  $F_X$  has the following properties:

- (i)  $F_X$  is non-decreasing
- (ii)  $\lim_{x \rightarrow -\infty} F_X(x) = 0$  and  $\lim_{x \rightarrow \infty} F_X(x) = 1$
- (iii)  $F_X$  is right-continuous

**Definition 34.2** (37.3). A random variable  $X$  is called discrete if there exists a countable set  $S \subseteq \mathbb{R}$  such that  $P(X \in S) = 1$ .

**Definition 34.3** (37.4). A random variable  $X$  is called continuous if its distribution  $\mu_X$  is absolutely continuous with respect to Lebesgue measure. In this case, the Radon-Nikodym derivative  $f_X = \frac{d\mu_X}{dx}$  is called the probability density function (PDF) of  $X$ .

**Definition 34.4** (37.5). *The characteristic function of a random variable  $X$  is defined by*

$$\phi_X(t) = \mathbb{E}[e^{itX}] = \int_{\mathbb{R}} e^{itx} d\mu_X(x).$$

**Theorem 34.1** (37.6: Inversion theorem for characteristic functions). *If  $\int_{\mathbb{R}} |\phi_X(t)| dt < \infty$ , then  $X$  has a continuous bounded density given by*

$$f_X(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \phi_X(t) dt.$$

**Theorem 34.2** (37.7: Uniqueness theorem for characteristic functions). *If two random variables have the same characteristic function, then they have the same distribution.*

## 35 Conditional expectation

**Definition 35.1** (38.1). *Let  $(\Omega, \mathcal{F}, P)$  be a probability space and  $\mathcal{G} \subseteq \mathcal{F}$  be a sub- $\sigma$ -algebra. The conditional expectation of an integrable random variable  $X$  given  $\mathcal{G}$  is a  $\mathcal{G}$ -measurable random variable  $\mathbb{E}[X|\mathcal{G}]$  such that for all  $G \in \mathcal{G}$ ,*

$$\int_G \mathbb{E}[X|\mathcal{G}] dP = \int_G X dP.$$

**Theorem 35.1** (38.2: Existence and uniqueness of conditional expectation). *For any integrable random variable  $X$  and any sub- $\sigma$ -algebra  $\mathcal{G}$ , the conditional expectation  $\mathbb{E}[X|\mathcal{G}]$  exists and is unique almost surely.*

**Proposition 35.1** (38.3: Properties of conditional expectation). *Let  $X, Y$  be integrable random variables and  $\mathcal{G}, \mathcal{H}$  be sub- $\sigma$ -algebras.*

- (i) *Linearity:  $\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}]$*
- (ii) *Tower property: If  $\mathcal{H} \subseteq \mathcal{G}$ , then  $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$*
- (iii) *If  $X$  is  $\mathcal{G}$ -measurable, then  $\mathbb{E}[X|\mathcal{G}] = X$*
- (iv) *If  $X$  is independent of  $\mathcal{G}$ , then  $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$*
- (v) *Monotonicity: If  $X \leq Y$ , then  $\mathbb{E}[X|\mathcal{G}] \leq \mathbb{E}[Y|\mathcal{G}]$*

**Theorem 35.2** (38.4: Conditional Jensen's inequality). *If  $\varphi$  is convex and  $X, \varphi(X)$  are integrable, then*

$$\varphi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\varphi(X)|\mathcal{G}].$$

**Definition 35.2** (38.5). *The conditional probability of an event  $A \in \mathcal{F}$  given  $\mathcal{G}$  is defined by*

$$P(A|\mathcal{G}) = \mathbb{E}[1_A|\mathcal{G}].$$