Chapter 0: Introduction

- **What is Analysis?**
- **What about Analysis?**
- **Why do Analysis?**

Background of mathematics

- **Number Systems**

  $\mathbb{N}$ = the set of **natural numbers**
  $\mathbb{Z}$ = the set of integers
  $\mathbb{Q}$ = the set of rational numbers
  $\mathbb{R}$ = the set of real numbers
  $\mathbb{C}$ = the set of complex numbers

- **Elementary algebra** $\to$ Advanced algebra $\to \cdots$
- **Euclidean geometry** $\to$ Analytic geometry $\to \cdots$
- **Calculus**

  $*$ What is the **limit**?
  $*$ What is the **continuity**?
  $*$ What is the **differential**?
  $*$ What is the **integral**?
  $*$ When is f **integrable** on $[a, b]$?
  $*$ What is the **fundamental theorem of calculus**?

Development of Analysis.

I. The First Wing

- **I. Newton** (1642-1727):
  - $*$ The method of fluxions (1666)
  - $*$ The sine and cosine series (1669)
- **G. W. Leibniz** (1646-1716)
  - $*$ The transmutation theorem (1673)
  - $*$ Leibniz series (1674)
- **The Bernoullis**
  - $*$ **Jakob** (1654-1705) and his figurate series (1689)
  - $*$ **Johann** (1667-1748) and the function x^x (1697)
- **L. Euler** (1707-1783)
  - $*$ The interpolation of $f(n) = n!$ (1729)
  - $*$ The resolution of **Jakob's challenge** (1748)
  - $*$ Differential $\frac{0}{0}$ and $d\sin x = \cos x dx$ (1755)
  - $*$ The integral $\int_0^1 \frac{\sin(\ln x)}{\ln x}\, dx$ (1768)

II. The Classical Wing

- **J. D'Alembert** (1717-1783):
  - ∗ From "limits" to differentials
  - ∗ D'Alembert's criterion for series
  - ∗ General solution of wave equation (1747)
- **J. Lagrange** (1736-1813):
  - ∗ From power series to differentials (1797)
  - ∗ **Applications of inequalities**
- **J. Fourier** (1768-1830):
  - ∗ Concept of functions
  - ∗ **Trigonometric series** (1807)
- **A. Cauchy** (1789-1857):
  - ∗ Limits, continuity, derivatives
  - ∗ The intermediate value theorem
  - ∗ The mean value theorem, integrals
  - ∗ **The fundamental theorem of calculus**
  - ∗ Cauchy's criterion for series
- **P. J. L. Dirichlet** (1805-1859):
  - ∗ Dirichlet function (1829),
  - ∗ Trigonometric series
  - ∗ Analytic number theory
- **G. F. B. Riemann** (1826-1866):
  - ∗ Riemann integral (1854)
  - ∗ Riemann function
  - ∗ Riemann rearrangement theorem (for series)

- **K. Weierstrass** (1815-1897):
  - “$\epsilon - \delta$” for limits
  - Uniform continuity of functions (**E. Heine**)
  - Uniform convergence of series (1841, 1894)
  - Weierstrass approximation theorem
  - Weierstrass M-test for series
  - **Weierstrass function** (P. Bois-Reymond, 1875)
- **G. Darboux** (1842-1917):
  - Simplification of Riemann integral
  - The Darboux theorem

Remark:

(i) How discontinuous can an integrable function be?

(ii) How discontinuous can a derivative be?

(iii) How, if at all, can we correct the deficiencies of the Riemann integral?

(iv) Recall sufficient conditions of the **fundamental theorem of calculus**

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

- F' is continuous on $[a, b]$ (**Cauchy**)
- F' is integrable on $[a, b]$ (**Darboux**)

III. The Modern Wing

- **G. Cantor** (1845-1918):
  - Completeness of the real numbers (1874)
  - Continuum Hypothesis
- **R. Baire** (1874-1932):
  - Nowhere-dense sets (due to **H. Hankel**)
  - The Baire category theorem (1899)
  - The Baire functions (1899)
- **H. Lebesgue** (1875-1941):
  - Measure of sets
  - Measurable functions
  - The Lebesgue integrals (1902)

The **main results** are

Theorem 1. A bounded function on $[a, b]$ is Riemann integrable **if and only if** it is continuous almost everywhere.

Theorem 2. If F is differential on $[a, b]$ with bounded derivative, then F' is Lebesgue integrable and

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

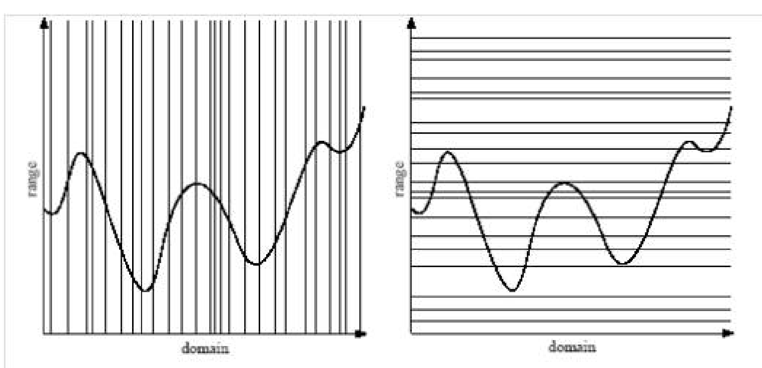
Henri Lebesgue

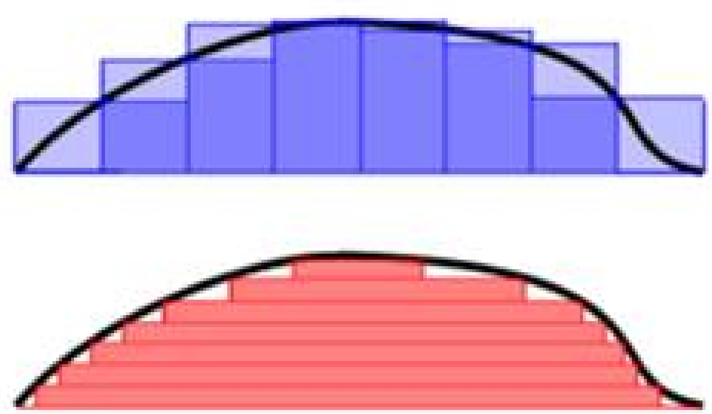
Henri Lebesgue (June 28, 1875-July 26, 1941) was a French mathematician most famous for his theory of integration, which was a generalization of the seventeenth century concept of integration summing the area between an axis and the curve of a function defined for that axis. His theory was published originally in his dissertation at the University of Nancy during 1902.

The Lebesgue Integral

• Riemann integral: partition the domain of the function f and obtain the Riemann sum $\sum_{j=1}^{n} f(\xi_j)|I_j|$.

• Lebesgue integral: partition the range of the function f and obtain the Lebesgue sum $\sum_{j=1}^{n} \eta_j |E_j|$.





• limits of functions → continuous functions → Riemann integrals

• measures of sets → measurable functions → Lebesgue integrals

Outline of Real Analysis.

- Sets
- Measures of sets
- Measurable functions
- Lebesgue integrals

References.

- E. M. Stein and R. Shakarchi, Real Analysis.
- J. N. McDonald and N.A. Weiss, **A Course in Real Analysis**.
- Zhou Min-Qiang, Functions of Real Variable (C).
- T. Tao, Real Analysis.
- P. R. Halmos, Measure theory.

System of natural numbers ($\mathbb{N}$)

- **What are natural numbers?**
- **What are operations of addition "+" and multiplication "×"?**

− −

We start with the following two fundamental concepts

⋆ the **zero number** "0"

⋆ the **increment operation** "++"

and define

$$
\begin{aligned}
1 &:= 0++,\\
2 &:= (0++)++,\\
3 &:= ((0++)++)++,\\
&\cdots\cdots\\
n &:= (\cdots((0++)++)++\cdots)++.
\end{aligned}
$$

Axiom I. 0 is a natural number.

Axiom II. The increment of a natural number is a natural number as well.

Axiom III. 0 is not the increment of any natural number.

Axiom IV. Different natural numbers must have different increments.

Axiom V. (Principle of induction) Let $p(n)$ be any property pertaining to natural numbers n. Suppose that $p(0)$ is true, and suppose that whenever $p(n)$ is true, $p(n++)$ is also true. Then $p(n)$ is true for every natural number n.

Proposition. 2 is a natural number.

Proof. By definition, $2 = 1{+}{+} = (0{+}{+}){+}{+}$. By **Axiom I**, 0 is a natural number; so is 2 by using **Axiom II** twice. ☐

Proposition. 3 is not equal to 0.

Proof. By definition, $3 = 2 + +$. Since we have proved 2 is a natural number (by **Axiom I** and **Axiom II**). Thus, the increment of the natural number 2 is not equal to 0 (**Axiom III**), i.e., $3 \neq 0$. ☐

Proposition. 3 is not equal to 2.

Proof. First, we can show that all the signs $0, 1, 2$, and 3 are natural numbers by using **Axioms I** and **II**.

Secondly, by definition, $3 = 2 + +$ and $2 = 1 + +$. Thus, by **Axiom IV**, it suffices to show that $2 \neq 1$. Also, by definition we have $2 = 1 + +$ and $1 = 0 + +$, and hence, it suffices to show that $1 \neq 0$ by **Axiom IV** again.

Finally, note that 1 is the increment of the natural number 0. Thus, $1 \neq 0$ due to **Axiom III**. ☐

Remark. Note how all of **Axioms I, II, III** and **IV** had to be used in above proofs. What about **Axiom V**?

Addition "+" of natural numbers

In order to define the operations of addition and multiplication, we need the following concepts

$\star$ the **set** X and $x \in X$, in particular, write

$$\mathbb{N} := \{0, 1, 2, 3, \cdots\},$$

$\star$ the **function** $f : \mathbb{N} \to \mathbb{N}$,
and the following result.

Proposition. (Recursive definition) Suppose that for every $n \in \mathbb{N}$, there is a function $f_n : \mathbb{N} \to \mathbb{N}$. Let c be a natural number. Then for every $n \in \mathbb{N}$, there is a unique $a_n \in \mathbb{N}$ such that $a_0 = c$ and $a_{n++} = f_n(a_n)$.

Proof. Define $a_0 := c$ and

$$a_{n++} := f_n(a_n), \quad n \in \mathbb{N}. \tag{0.1}$$

First, note that $0 \neq n{+}{+}$ for all $n \in \mathbb{N}$ (by using **Axiom III**). Thus, none of the definitions $a_{n++} := f_n(a_n)$, given in the procedure (0.1), will redefine the value of a_0. This implies that the value of a_0 is uniquely determined.

Suppose inductively that the value of a_n is uniquely determined. Then by the definition of function, the value of $f_n(a_n)$ is uniquely determined as well. In addition, note that $m++ \neq n++$ for $m \neq n$ (**Axiom IV**). This implies that none of the other definitions $a_{m++} := f_m(a_m)$ will redefine the value of a_{n++}, that is, the value of a_{n++} be defined only once. Thus, the value of a_{n++} is uniquely determined.

So, by the principle of induction (**Axiom V**), the value of a_n is uniquely determined for every $n \in \mathbb{N}$. $\square$

By using the recursive definition, we can give the definition of addition "+" as follows.

Definition. (Addition) Let $m, n \in \mathbb{N}$. Define $0 + m := m$. Suppose that $n + m$ is well defined, we then define $(n++) + m := (n + m) + +$.

Remark. By the last proposition (recursive definition), we observe that, for every $n \in \mathbb{N}$, the sum $n+m$ is uniquely determined. Indeed, note that f given by

$$f(n) := n + +, \quad n \in \mathbb{N}$$

is a function from $\mathbb{N}$ to $\mathbb{N}$ (WHY). By using the signs in the last proposition ($f_n = f, a_n = n + m$ for all $n \in \mathbb{N}$), our recursive definition is given by $a_0 := m$ and

$$a_{n++} := (n + m) + + = f(n + m) = f_n(a_n).$$

Thus, all the values for $a_0, a_1, a_2, \cdots$ are uniquely determined. That is, **the addition of natural numbers is well defined**.

Example. Let $n \in \mathbb{N}$. Then

$$\begin{aligned} 1 + n &= n + +, \\ 2 + n &= (n + +) + +, \\ 3 + n &= ((n + +) + +) + +. \end{aligned}$$

Proposition. Let $m, n \in \mathbb{N}$. Then $n + m \in \mathbb{N}$.

Proof. Fix m and apply induction to n. Note that $0 \in \mathbb{N}$ (**Axiom I**) and $0 + m = m \in \mathbb{N}$. Suppose inductively that $k + m \in \mathbb{N}$ for some $k \in \mathbb{N}$. Then

$$(k + +) + m = (k + m) + + \in \mathbb{N}$$

by **Axiom II**. Thus, $n+m \in \mathbb{N}$ for all $n \in \mathbb{N}$ due to **Axiom V**. □

Lemma 1. Let $m \in \mathbb{N}$. Then $m + 0 = m$.

Proof. Clearly, $0+0=0$. Suppose inductively that $k+0=k$. Then by the definition of addition,

$$(k++)+0=(k+0)++=k++.$$

Thus, $m+0=m$ for all $m\in\mathbb{N}$ by the principle of induction. □

Lemma 2. Let $m,n\in\mathbb{N}$. Then

$$n+(m++)=(n+m)++.$$

Proof. Fix m and apply the principle of induction on n. Indeed, it is clear that (by the first step of definition of addition)

$$0+(m++)=m++=(0+m)++.$$

Suppose inductively that $k+(m++)=(k+m)++$. Then

$$\begin{aligned}&(k++)+(m++)\\=&(k+(m++))++\\=&((k+m)++)++\\=&((k++)+m)++.\end{aligned}$$

Thus, we obtain the desired conclusion by the principle of induction. □

Corollary. Let $m\in\mathbb{N}$. Then $m++=m+1=1+m$.

Proof. The first equality is a direct consequence of Lemmas 2 and 1. Indeed,

$$m+1=m+(0++)=(m+0)++=m++.$$

□

Further on, by using the principle of induction, we can show that the natural numbers obey so-called commutative, associative and cancellation laws.

Proposition. (Commutative law) Let $a, b \in \mathbb{N}$. Then

$$a + b = b + a.$$

Proof. Fix a and apply induction on b. Clearly,

$$a + 0 = 0 = 0 + a.$$

Suppose inductively that $a + k = k + a$. Then

$$\begin{aligned} a + (k++) =& (a + k) ++ \\ =& (k + a) ++ = (k++) + a. \end{aligned}$$

Thus, we obtain the desired conclusion by the principle of induction. □

Proposition. (Associative law) Let $a, b, c \in \mathbb{N}$. Then

$$(a + b) + c = a + (b + c).$$

Proof. Fix b, c and apply induction on a. Clearly,

$$(0 + b) + c = b + c = 0 + (b + c).$$

Suppose inductively that $(k + b) + c = k + (b + c)$. Then

$$\begin{aligned} ((k++) + b) + c =& ((k + b)++) + c \\ =& ((k + b) + c)++ \\ =& (k + (b + c))++ = (k++) + (b + c). \end{aligned}$$

Thus, we obtain the desired conclusion by the principle of induction. □

Proposition. (Cancellation law) Let $a, b, c \in \mathbb{N}$. If $a + b = a + c$ then $b = c$.

Proof. Fix b, c and apply induction on a. Clearly, $0 + b = 0 + c$ implies that $b = c$.

Suppose inductively that whenever $k + b = k + c$ we have $b = c$. Let $(k++) + b = (k++) + c$. Then

$$\begin{aligned}(k + b) ++ =& (k++) + b \\ =& (k++) + c = (k + c)++,\end{aligned}$$

so that $k + b = k + c$ (by **Axiom IV**), and hence, the inductive hypothesis implies that $b = c$.

Thus, we obtain the desired conclusion by the principle of induction. □

We finally introduce the concept of positive natural number which will be used to construct so-called order of natural numbers in the next subsection.

Definition. A natural number n is said to be **positive** if n is not equal to 0.

Proposition. Let $m, n \in \mathbb{N}$. If m is positive, then $m+n$ is positive as well.

Proof. Let m be positive. We apply principle of induction on n. It is clear that $m + 0 = m$ is positive.

Suppose inductively that $m + k$ is positive. Then, by the associative law of addition, we have

$$m + (k + 1) = (m + k) + 1 = (m + k)++.$$

This implies that $m+(k+1) \neq 0$ (WHY). Thus, $m+(k+1)$ is positive by the definition of positivity of natural numbers.

Finally, by the principle of induction, we obtain the desired conclusion. □

Corollary. Let $a, b \in \mathbb{N}$. If $a + b = 0$, then $a = 0$ and $b = 0$.

Order of natural numbers

$$\geq, \quad =, \quad \leq, \quad >, \quad <$$

Definition. (Order of natural numbers) Let $m, n \in \mathbb{N}$. We say that n is greater than or equal to m (write $n \geq m$ or $m \leq n$) if $n = m + a$ for some $a \in \mathbb{N}$, and we say n is strictly greater than m (write $n > m$ or $m < n$) if $n \geq m$ and $n \neq m$.

Proposition. (Basic properties of order) Let $a, b, c \in \mathbb{N}$. The following statements hold.

(i) $a \geq a$.

(ii) If $a \geq b$ and $b \geq c$, then $a \geq c$.

(iii) If $a \geq b$ and $b \geq a$, then $a = b$.

(iv) $a \geq b$ if and only if $a + c \geq b + c$ for all/some $c \in \mathbb{N}$.

(v) $a < b$ if and only if there is a positive $d \in \mathbb{N}$ such that $b = a + d$, or equivalently, $a{+}{+} \leq b$.

Proposition. (Trichotomy of order) Let $a, b \in \mathbb{N}$. Then exactly one of the following statements is true:

$$a > b, \quad a = b, \quad a < b.$$

Proof. Step I. It can not be more than one of the statements holding at the same time (**by the method of contradiction**).

Step II. There holds at least one of the statements (**by the method of induction** with b fixed).

Indeed, for $a = 0$, it is clear that $a = b$ if $b = 0$. If $b \neq 0$, from the equality $b = a + b$ we observe that $b > a$. Thus, there holds at least one of above three statements.

For $n = k$, suppose inductively that there holds at least one of the statements:

$$k > b, \quad k = b, \quad k < b.$$

Then, for $n = k + 1$, we have

(i) if $k > b$, then $k + 1 > k > b$;

(ii) if $k = b$, then $k + 1 > k = b$;

(iii) if $k < b$, then $k + 1 \leq b$ (WHY), so that $k + 1 = b$ or $k + 1 < b$.

Thus, there holds at least one of the statements:

$$k + 1 > b, \quad k + 1 = b, \quad k + 1 < b.$$

Finally, by the principle of induction, we obtain the desired conclusion. $\square$

Multiplication "×" of natural numbers

Definition. **(Multiplication)** Let $m, n \in \mathbb{N}$. Define $0 \times m := 0$. Suppose that $n \times m$ is well defined, we then define $(n++) \times m := (n \times m) + m$.

Remark. Observe that the multiplication of natural numbers is well defined. Indeed, by letting $a_n = n \times m$ and $f_n = f$ (the signs in the proposition of recursive definition) with $f : \mathbb{N} \to \mathbb{N}$ defined by

$$f(n) := n + m, \quad n \in \mathbb{N},$$

our multiplication is given by $a_0 := 0$ and

$$a_{n++} := (n \times m) + m = f(a_n), \quad n \in \mathbb{N},$$

Note that f is a function since the addition of natural numbers is well defined. Thus, the multiplication of natural numbers is also well defined by the proposition of recursive definition.

Example. Let $n \in \mathbb{N}$. Then

$$\begin{aligned} 1 \times n &= n, \\ 2 \times n &= n + n, \\ 3 \times n &= n + n + n. \end{aligned}$$

This implies that, just as addition is iterated incrementation, multiplication is incremented addition.

Proposition. Let $m, n \in \mathbb{N}$. Then $n \times m \in \mathbb{N}$.

Proof. Fix m and apply induction (**Axiom V**) on n.

First, by the definition $0 \times m = m \in \mathbb{N}$.

Second, suppose that $k \times m \in \mathbb{N}$. Note that

$$(k++) \times m = (k \times m) + m \quad \text{(Definition)}.$$

Since the set of natural numbers $\mathbb{N}$ is closed under the operation of addition, $(k\times m)+m \in \mathbb{N}$, and hence, $(k{+}{+})\times m \in \mathbb{N}$. Thus, $n \times m \in \mathbb{N}$ for every $n \in \mathbb{N}$ by **Axiom V**.

Finally, the arbitrary of m implies that $n \times m \in \mathbb{N}$ for all $m, n \in \mathbb{N}$. □

Proposition. (Commutative law) Let $m, n \in \mathbb{N}$. Then

$$n \times m = m \times n.$$

Proof. Step I. Prove that $m \times 0 = 0$ for every $m \in \mathbb{N}$.

Step II. Prove that

$$n \times (m++) = (n \times m) + n, \quad m, n \in \mathbb{N}.$$

Step III. Prove the desired conclusion. □

Remark. We write $a \times b := ab$ for short and use convention that multiplication takes precedence over addition, for example, $ab + c = (a \times b) + c$.

Proposition. Let $a, b \in \mathbb{N}$. The following statements hold.

(i) If a, b are positive, then ab is also positive.

(ii) $ab = 0$ if and only if $a = 0$ or $b = 0$.

(iii) $ab \neq 0$ if and only if $a \neq 0$ and $b \neq 0$.

Proof. (i) Let a, b be positive. Since $a \neq 0$, there is a natural number a' such that $a' + + = a$ (Definition). By the definition of multiplication, we have

$$a \times b = (a' + +) \times b = (a' \times b) + b,$$

and the last term is positive due to the positivity of b. This proves the positivity of ab.

(ii) ($\Rightarrow$) Let $ab = 0$. Assume $a \neq 0$ and $b \neq 0$. It follows from (i) that $ab \neq 0$ which contradicts $ab = 0$. Thus, we have $a = 0$ or $b = 0$.

($\Leftarrow$) It is easy to verify that $0 \times m = m \times 0 = 0$, which has been given in the proof of the last proposition.

Finally, it is clear that (ii) $\Leftrightarrow$ (iii). $\square$

Proposition. (Distributive law) Let $a, b, c \in \mathbb{N}$. Then

$$a(b + c) = ab + ac,$$
$$(b + c)a = ba + ca.$$

Proof. Fix b, c and apply induction on a. It is clear that

$$0(b + c) = 0 = 0b + 0c.$$

Suppose inductively that $k(b + c) = kb + kc$. Then by the definition of multiplication, we have

$$\begin{aligned}(k++)(b + c) =& k(b + c) + (b + c)\\ =&(kb + b) + (kc + c) = (k++)b + (k++)c.\end{aligned}$$

Thus, we obtain the desired conclusion by the principle of induction. $\square$

Proposition. (Associative law) Let $a, b, c \in \mathbb{N}$. Then

$$(a \times b) \times c = a \times (b \times c).$$

Proposition. (Multiplication preserves order) Let $a, b \in \mathbb{N}$ and $a < b$. If $c \in \mathbb{N}$ is positive, then $ac < bc$.

Proof. Fix $a < b$ and apply induction on c. It is clear that

$$a1 = a < b = b1.$$

Suppose inductively that $ak < bk$. Then, by the basics of addition of natural numbers, we have

$$\begin{aligned} a(k++) =& ak + a \\ <& bk + b = b(k++). \end{aligned}$$

Thus, we obtain the desired conclusion by the principle of induction. □

Corollary. **(Cancellation law)** Let $a, b, c \in \mathbb{N}$ such that $ac = bc$. If c is positive, then $a = b$.

Remark. Let $a, b, c \in \mathbb{N}$. By analogous argument, we can show that

(i) $ac < bc$ with c negative implies that $a > b$,

(ii) $ac = bc$ with c negative implies that $a = b$.

Proposition. **(Euclidean algorithm)** Let $n \in \mathbb{N}$ and let $q \in \mathbb{N}$ be positive. Then there exist two natural numbers m, r such that $0 \leq r < q$ and $n = mq + r$.

Proof. Fix q and apply induction on n. Clearly,

$$0 = 0 \times 0 + 0,$$

where $m = r = 0 \in \mathbb{N}$.

Suppose inductively that there are $m_k, r_k \in \mathbb{N}$ with $0 \leq r_k < q$ such that $k = m_k q + r_k$. Note that $1 \leq r_k + 1 \leq q$.

(i) If $r_k + 1 = q$, then

$$\begin{aligned} k + 1 =& m_k q + r_k + 1 \\ =& (m_k + 1)q + 0 = mq + r, \end{aligned}$$

with $m = m_k + 1$ and $r = 0$.

(ii) If $r_k + 1 < q$, then

$$k + 1 = m_k q + (r_k + 1) = mq + r,$$

with $m = m_k$ and $r = r_k + 1$.

Therefore, by the principle of induction, we obtain the desired conclusion. $\square$

Remark. Euclidean algorithm marks the beginning of **number theory**.

Finally, we define exponentiation for natural number exponents. Just like addition was recursive increment and multiplication was recursive addition, exponentiation is recursive multiplication.

Definition. (Exponentiation) Let $m \in \mathbb{N}$. To raise m to the power 0, we define $m^0 := 1$. Now suppose that recursively that m^n has been defined for some $n \in \mathbb{N}$, then we define $m^{n++} := m^n \times m$.

Example. Let $m \in \mathbb{N}$. Then

$$\begin{aligned} m^1 &= 1 \times m, \\ m^2 &= m \times m, \\ m^3 &= m \times m \times m. \end{aligned}$$

So far, we have constructed the system of natural numbers $(\mathbb{N}, +, \times)$. Recall the expansion of system of numbers

$$\mathbb{N} \to \mathbb{Z} \to \mathbb{Q} \to \cdots .$$

Problem. Could you give appropriate definition of systems of **integers** and **rational numbers**:

$$(\mathbb{Z}, +, -, \times), \quad (\mathbb{Q}, +, -, \times, \div)?$$

Problem. What about the system of **real numbers**?

Mathematical logic

In order to extend the system of natural numbers to that of integers, we need some basics of mathematical logic:

I. Mathematical statement.

II. Logical connectives.

III. Equality (=)**.**

IV. The structure of proofs.

I. Mathematical statement is a judgement concerning various mathematical objects (numbers, functions, etc.) and relations between them (addition, multiplications, etc.), which obeys the following basic axiom of mathematical logic: **a statement is either true or false, but not both**.

II. One can make a **compound statement** from more primitive statements by using **logical connectives** such as **and**, **or**, **not**, **if-then** ($\Rightarrow$), **if-and-only-if** ($\Leftrightarrow$), and so forth.

III. Equality (denoted by "=") is a relation linking two objects x, y of the same type (e.g., two natural numbers) which obeys the following four **axioms of equality**:

- (**Reflexive axiom**) $x = x$.
- (**Symmetry axiom**) If $x = y$, then $y = x$.
- (**Transitive axiom**) If $x = y$ and $y = z$, then $x = z$.
- (**Substitution axiom**) If $x = y$, then $f(x) = f(y)$ for all operations or functions f. Similarly, for any property $P(x)$ depending on x, if $x = y$, then $P(x)$ and $P(y)$ are equivalent statements.

IV. Prove the proposition "$A \Rightarrow B$".

Proof 1. Suppose that A is true. Since A is true, C is true. Since C is true, D is true. Since D is true, B is true. This completes the proof. □

Proof 2. To show B, it suffices to show D. Since C implies D, we just need to show C. But C follows from A. Thus, A implies B. □

Proof 3. To show that $\overline{B} \Rightarrow \overline{A}$ (WHY). □

Proof 4. To show that $\overline{B}$ implies something which is known to be false. That is, **assume** B **is false** and show that this implies some statement simultaneously true and not true. This contradicts the axiom on statements. □

The system of integers ($\mathbb{Z}$)

Definition. (Integers) Let $a, b \in \mathbb{N}$. An **integer** is an expression of the form $a \smile b$. We say that two integers $a \smile b$ and $c \smile d$ are **equal** if $a + d = c + b$. The set of all integers is denoted by $\mathbb{Z}$.

Definition. (Addition) The **sum of two integers**, $(a \smile b) + (c \smile d)$ is defined by

$$(a \smile b) + (c \smile d) := (a + c) \smile (b + d),$$

Definition. (Multiplication) The **product of two integers**, $(a \smile b) \times (c \smile d)$ is defined by

$$(a \smile b) \times (c \smile d) := (ac + bd) \smile (ad + bc),$$

Remark. The concepts of equality, addition and multiplication of integers are all well defined. Indeed, it suffices to verify the reflexivity, symmetry, transitivity and substitution axioms:

(i) **(Reflexivity)** $a \smile b = a \smile b$.

(ii) **(Symmetry)** $a \smile b = c \smile d \Rightarrow c \smile d = a \smile b$.

(iii) **(Transitivity)** $a \smile b = c \smile d$ and $c \smile d = e \smile f$ implies that $a \smile b = e \smile f$.

(iv) **(Substitution)** $a \smile b = a' \smile b'$ implies that

$$(a \smile b) + (c \smile d) = (a' \smile b') + (c \smile d),$$
$$(c \smile d) + (a \smile b) = (c \smile d) + (a' \smile b'),$$

and

$$(a \smile b) \times (c \smile d) = (a' \smile b') \times (c \smile d),$$
$$(c \smile d) \times (a \smile b) = (c \smile d) \times (a' \smile b').$$

Remark. The integers $n \smile 0$ behave in the same way as the natural numbers n, so that we may **identify** the natural numbers with integers by setting $n \equiv n \smile 0$. Indeed, this gives an isomorphism between the natural numbers n and the integers of the form $n \smile 0$, that is,

(i) $n \smile 0 = m \smile 0$ if and only if $n = m$,

(ii) $(n \smile 0) + (m \smile 0) = (n + m) \smile 0$,

(ii) $(n \smile 0) \times (m \smile 0) = (n \times m) \smile 0$.

Definition. (Negative integers) Let $a \smile b$ be an integer. We define the **negation** $-(a \smile b)$ to be the integer $(b \smile a)$.

Remark. For natural number $n = n \smile 0$, we have $-n = 0 \smile n$. Note that $n = 0 \Leftrightarrow -n = 0$.

Proposition. Let $m, n \in \mathbb{N}$. Then

$$(-n)m = n(-m) = -(nm).$$

Lemma. (Trichotomy of integers) Let $x \in \mathbb{Z}$. Then exactly one of the following three statements is true:

(i) x is zero,

(ii) x is equal to a positive natural number,

(iii) x is the negation of a positive natural number.

Proof. Let $x = a \smile b$ with $a, b \in \mathbb{N}$. By the trichotomy of natural numbers, exactly one of the following statements holds: $a = b, \quad a > b, \quad a < b$.

(i) If $a = b$, then $x = 0 \smile 0 = 0$ by the definition of integers.

(ii) If $a > b$, then there is a positive $b' \in \mathbb{N}$ such that $a = b + b'$. Thus, $x = b' \smile 0 = b'$.

(iii) If $a < b$, then there is a positive $a' \in \mathbb{N}$ such that $b = a + a'$. Thus, $x = 0 \smile a' = -a'$.

$\square$

Proposition. (Laws of algebra for integers) Let $x, y, z \in \mathbb{Z}$. Then we have

$$\begin{aligned} x + y =& y + x, \\ (x + y) + z =& x + (y + z), \\ x + 0 =& 0 + x = x, \\ x + (-x) =& (-x) + x = 0, \\ xy =& yx, \\ (xy)z =& x(yz), \\ x1 =& 1x = x, \\ x(y + z) =& xy + xz, \\ (y + z)x =& yx + zx. \end{aligned}$$

Definition. (Substraction) Let $x, y \in \mathbb{Z}$. The **difference** $x - y$ is defined by

$$x - y := x + (-y).$$

Proposition. Let $x, y \in \mathbb{N}$. Then $x - y = x \smile y$.

Remark. Observe that every integer is the difference of two natural numbers.

Proposition. Let $a \in \mathbb{Z}$. Then $a + (-a) = 0$.

Proposition. Let $a \in \mathbb{Z}$. Then $a = 0 \Leftrightarrow -a = 0$.

Proposition. Let $a, b \in \mathbb{Z}$. Then $a = b \Leftrightarrow a - b = 0$.

Proposition. Let $a, b \in \mathbb{Z}$. If $ab = 0$, then either $a = 0$ or $b = 0$.

Proposition. Let $a, b, c \in \mathbb{Z}$. Then $ab - ac = a(b - c)$.

Proposition. Let $a, b, c \in \mathbb{Z}$. If $ac = bc$ and if $c \neq 0$, then $a = b$.

Order of integers

Definition. (Order of integers) Let $a, b \in \mathbb{Z}$. We say that a is greater than or equal to b (write $a \geq b$ or $b \leq a$) if $a = b + n$ for some $n \in \mathbb{N}$, and we say a is strictly greater than b (write $a > b$ or $b < a$) if $a \geq b$ and $a \neq b$.

Proposition. (Basic properties of order) Let $a, b, c \in \mathbb{Z}$. The following statements hold.

(i) $a > b$ if and only if $a - b$ is a positive natural number.

(ii) If $a > b$, then $a + c > b + c$.

(iii) If $a > b$ and c is positive, then $ac > bc$.

(iv) If $a > b$, then $-a < -b$.

(v) If $a > b$ and $b > c$, then $a > c$.

(vi) Exactly one of the following three statements is true:

$$a > b, \quad a = b, \quad a < b.$$

Remark. Notice that the principle of induction is **not** valid for integers. Can you give an example to explain it?

Example. Fix $a, b \in \mathbb{N}$ and consider the statement $P(n)$: $a > b$ implies that $a(1 + n) > b(1 + n)$ for all $n \in X$.

(i) $X = \mathbb{N}$.

(ii) $X = \mathbb{Z}$.

The system of rational numbers ($\mathbb{Q}$)

Definition. (Rational numbers) Let $a, b \in \mathbb{Z}$ and $b \neq 0$. A **rational number** is an expression of the form $a//b$. We say that two rational numbers $a//b$ and $c//d$ are **equal** if $ad = cb$. The set of all rational numbers is denoted by $\mathbb{Q}$.

Definition. (Addition) The **sum of two rational numbers**, $(a//b) + (c//d)$ is defined by

$$(a//b) + (c//d) := (ad + cb)//(bd).$$

Definition. (Multiplication) The **product of two rational numbers**, $(a//b) \times (c//d)$ is defined by

$$(a//b) \times (c//d) := (ac)//(bd).$$

Definition. (Negation) The **negation of a rational number**, $-(a//b)$ is defined by

$$-(a//b) := (-a)//b.$$

Remark. Observe that rational numbers $a//1$ behave in the same way as integers a:

(i) $a//1 = b//1$ if and only if $a = b$,

(ii) $(a//1) + (b//1) = (a + b)//1$,

(iii) $(a//1) \times (b//1) = (a \times b)//1$,

(iv) $-(a//1) = (-a)//1$.

Thus, in the sequel, we will identify a with $a//1$ for every integer a, that is, $a \equiv a//1$.

Proposition. Let $x = a//b \in \mathbb{Q}$. Then $x = 0 \Leftrightarrow a = 0$.

Proof. Since $0 \in \mathbb{Z}$, we have $0 = 0//1$. By the definition of equality of rational number,

$$0//1 = a//b \Leftrightarrow 0b = a1 \Leftrightarrow a = 0.$$

This proves that desired conclusion. $\square$

Definition. **(Reciprocal)** Let $x = a//b$ is a non-zero rational number. The **reciprocal** of x (denoted by x^{-1}) is defined by

$$x^{-1} := b//a.$$

By using the negation and reciprocal, we can define the substraction and division of two rational numbers, respectively.

Definition. **(Substraction)** Let $x, y \in \mathbb{Q}$. The **difference** $x - y$ is defined by

$$x - y := x + (-y).$$

Definition. **(Division)** Let $x, y \in \mathbb{Q}$ and $y \neq 0$. The **quotient** x/y is defined by

$$x/y := x \times y^{-1}.$$

Remark. Observe that $a/b = a//b$ for all $a, b \in \mathbb{Z}$ with $b \neq 0$. Indeed, let $a, b \in \mathbb{Z}$ and $b \neq 0$. Then

$$a/b = a \times b^{-1} = (a//1) \times (1//b) = (a1)//(1b) = a//b.$$

Thus, in the sequel, we also use the more customary a/b instead of $a//b$.

Proposition. (Laws of algebra for rational numbers) Let $x, y, z \in \mathbb{Q}$. Then we have

$$\begin{aligned} x + y =& y + x, \\ (x+y)+z =& x + (y+z), \\ x + 0 =& 0 + x = x, \\ x + (-x) =& (-x) + x = 0, \\ xy =& yx, \\ (xy)z =& x(yz), \\ x1 =& 1x = x, \\ x(y+z) =& xy + xz, \\ (y+z)x =& yx + zx. \end{aligned}$$

If further $x \neq 0$, then

$$xx^{-1} = x^{-1}x = 1.$$

Definition. A rational number x is said to be **positive** if $x = a/b$ for some positive integers a and b, and x is said to be **negative** if $x = -y$ for some positive rational number y.

Lemma. Let $x, y \in \mathbb{Q}$. Then

(i) $x + (-x) = 0$.

(ii) $x = y \Leftrightarrow -x = -y$;

(iii) x is positive if and only if $-x$ is negative.

Proof. (ii) The necessary is obvious. Let $-x = -y$. By taking $x = a//b$ and $y = c//d$, we have

$$(-a)//b = -(a//b) = -(c//d) = (-c)//d.$$

Thus, $(-a)d = (-c)b$, which implies that $ad = cb$. This proves $a//b = c//d$.

(iii) is a direct consequence of the statement (ii). $\square$

Proposition. (Trichotomy of rational numbers) Let $x \in \mathbb{Q}$. Then exactly one of the following three statements is true:

(i) x is equal to 0,

(ii) x is a positive rational number,

(iii) x is a negative rational number.

Proof. Let $x = a//b$, where $a, b \in \mathbb{Z}$ and $b \neq 0$. By the trichotomy of integers, we have a is zero, or positive natural number, or negative natural number; and have b is positive natural number or negative natural number. Clearly, it can not be more than one of the three statements holding at the same time.

(i) If $a = 0$, then $x = 0$ (no matter $b > 0$ or $b < 0$);

(ii) If a and b are both positive, then x is positive by the definition of positivity of rational number, directly.

(iii) If a is positive and b is negative, then

$$a//b = a//(-(-b)) = -(a//(-b)).$$

Thus, $x = a//b$ is negative. So is the case that a is negative and b is positive.

(iv) If a and b are both negative, then

$$a//b = (-a)//(-b).$$

Thus, $x = a//b$ is positive. □

Definition. (Order of rational numbers) Let $x, y \in \mathbb{Q}$. We say that $x > y$ if $x - y$ is a positive rational number, and $x < y$ if $x - y$ is a negative rational number. In addition, we write $x \geq y$ if either $x > y$ or $x = y$, and $x \leq y$ if either $x < y$ or $x = y$.

Remark. Let $x \in \mathbb{Q}$. Since $x - 0 = 0$, it is clear that $x > 0$ if and only if x is positive. Also, $x < 0$ if and only if x is negative.

Proposition. (Basic properties of order) Let $x, y, z \in \mathbb{Q}$. The following statements hold.

(i) Exactly one of the following three statements is true:

$$x > y, \quad x = y, \quad x < y.$$

(ii) $x > y$ if and only if $y < x$.

(iii) If $x > y$ and $y > z$, then $x > z$.

(iv) If $x > y$, then $x + z > y + z$.

(v) If $x > y$ and z is positive, then $xz > yz$.

(vi) If $x > y$ and z is negative, then $xz < yz$.

So far, we have constructed the system $(\mathbb{Q}, +, -, \times, \div, \leq)$ which has been proved to be an **ordered field**. Finally, we introduce two particularly useful ones: "absolute value" and "distance" which will be used to construct the system of so-called "real numbers".

Absolute value and distance of rational numbers

Definition. (Absolute value) Let $x \in \mathbb{Q}$. The **absolute value** of x (denoted by $|x|$) is defined as follows:

(i) If x is zero, then $|x| := 0$.

(ii) If x is positive, then $|x| := x$.

(iii) If x is negative, then $|x| := -x$.

Remark. It is clear that $\pm x \leq |x|$ for every $x \in \mathbb{Q}$.

Definition. (Distance) Let $x, y \in \mathbb{Q}$. The quantity $|x - y|$ is called the **distance** between x and y. We also write $d(x, y) := |x - y|$ for convenience.

Proposition. (Basic properties of absolute value) Let $x, y, z \in \mathbb{Q}$. The following statements hold.

(i) **(Non-degeneracy)** $0 \leq |x|$. Also, $|x| = 0$ if and only if $x = 0$.

(ii) **(Triangle inequality)** $|x + y| \leq |x| + |y|$.

(iii) $-y \leq x \leq y$ if and only if $|x| \leq y$.

(iv) $|xy| = |x||y|$. In particular, $|-x| = |x|$.

Proposition. (Basic properties of distance) Let $x, y, z \in \mathbb{Q}$. The following statements hold.

(i) **(Non-degeneracy)** $0 \leq d(x, y)$. Also, $d(x, y) = 0$ if and only if $x = y$.

(ii) **(Symmetry)** $d(x, y) = d(y, x)$.

(iii) **(Triangle inequality)** $d(x, z) \leq d(x, y) + d(y, z)$.

Finally, it is necessary to point out that

Proposition. Let $x \in \mathbb{Q}$. Then there is a unique $n \in \mathbb{Z}$ such that $n \leq x < n + 1$.

Proof. (Existence) Without loss of generality, we can take $x = a//b$ with $a, b \in \mathbb{Z}$ and $b > 0$ (WHY).

(i) If $a \geq 0$, then, by the Euclidean algorithm, there are two numbers $n, r \in \mathbb{N}$ with $0 \leq r < b$ such that $a = nb + r$, so that $0 \leq a - nb < b$, or, equivalently,

$$nb \leq a < (n+1)b.$$

This implies that $n \leq x < n+1$ (WHY).

(ii) If $a < 0$, then $-x > 0$. Thus, we can complete the proof of existence by using the step (i) (HOW).

(Uniqueness) Suppose that $m, n \in \mathbb{Z}$ satisfying

$$m \leq x < m+1, \quad n \leq x < n+1.$$

Then

$$m \leq x < n+1, \quad n \leq x < m+1,$$

and hence, $-1 < m - n < 1$. Thus, $m - n = 0$. □

Proposition. Let $x, y \in \mathbb{Q}$ with $x < y$. Then there is a $z \in \mathbb{Q}$ such that $x < z < y$.

Proof. Write $z := (x+y)/2$. Note that $x < y$ implies that

$$x/2 < y/2. \quad \text{(WHY)}$$

Thus,

$$z = (x/2) + (y/2) < (y/2) + (y/2) = y,$$

and

$$x = (x/2) + (x/2) < (y/2) + (x/2) = z.$$

□

The set of real numbers ($\mathbb{R}$)

Definition. (Sequence) Let $m \in \mathbb{Z}$. A **sequence** $\{a_n\}_{n=m}^{\infty}$ of rational numbers is any function from the set $X := \{n \in \mathbb{Z} : n \geq m\}$ to $\mathbb{Q}$. More informally, $\{a_n\}_{n=m}^{\infty}$ is a collection of rational numbers

$$a_m, a_{m+1}, a_{m+2}, \cdots .$$

In addition, if $a_n = c \in \mathbb{Q}$ for all $n \in X$, we write

$$\{c\}_{n=m}^{\infty} := \{a_n\}_{n=m}^{\infty}$$

for short.

Definition. (Cauchy sequence of rational numbers) A sequence $\{a_n\}_{n=1}^{\infty}$ of rational numbers is said to be a **Cauchy sequence** if for every rational number $\epsilon > 0$, there is a natural number N such that

$$|a_m - a_n| < \epsilon$$

for all $m, n \geq N$.

Definition. (Equality of Cauchy sequences) Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be two Cauchy sequences. We say that $\{a_n\}_{n=1}^{\infty}$ is **equal** to $\{b_n\}_{n=1}^{\infty}$ (denoted by $\{a_n\}_{n=1}^{\infty} = \{b_n\}_{n=1}^{\infty}$) if for every rational number $\epsilon > 0$, there is a natural number N such that

$$|a_n - b_n| < \epsilon$$

for all $n \geq N$.

Lemma. (Subsequence of Cauchy sequences) Every of subsequence $\{a_{n_k}\}_{k=1}^{\infty}$ of a Cauchy sequence $\{a_n\}_{n=1}^{\infty}$ is also a Cauchy sequence. In addition, $\{a_{n_k}\}_{k=1}^{\infty}$ is equal to $\{a_k\}_{k=1}^{\infty}$ itself.

Lemma. (Boundedness of Cauchy sequences) A Cauchy sequence is bounded, that is to say, there exists a rational number M such that $|a_n| \leq M$ for all $n \in \mathbb{N}$.

Lemma. (Closedness for addition and multiplication of Cauchy sequences) Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be two Cauchy sequences. Then $\{a_n + b_n\}_{n=1}^{\infty}$ and $\{a_n b_n\}_{n=1}^{\infty}$ are Cauchy sequences as well.

Lemma. Let $\{a_n\}_{n=1}^{\infty}, \{a'_n\}_{n=1}^{\infty}, \{b_n\}_{n=1}^{\infty}$ and $\{b'_n\}_{n=1}^{\infty}$ be four Cauchy sequences. If

$$\{a_n\}_{n=1}^{\infty} = \{a'_n\}_{n=1}^{\infty}, \quad \{b_n\}_{n=1}^{\infty} = \{b'_n\}_{n=1}^{\infty},$$

then $\{a_n+b_n\}_{n=1}^{\infty} = \{a'_n+b'_n\}_{n=1}^{\infty}$ and $\{a_n b_n\}_{n=1}^{\infty} = \{a'_n b'_n\}_{n=1}^{\infty}$.

Definition. (Real numbers) We call a Cauchy sequence a **real number**, and define two equivalent Cauchy sequences as the same real numbers. The set of all real numbers is denoted by $\mathbb{R}$.

Definition. (Addition of real numbers) Let $a := \{a_n\}_{n=1}^{\infty}$ and $b := \{b_n\}_{n=1}^{\infty}$ be two Cauchy sequences. The **sum** of a and b (denoted by $a + b$) is defined by

$$a + b := \{a_n + b_n\}_{n=1}^{\infty}.$$

Definition. (Multiplication of real numbers) Let $a := \{a_n\}_{n=1}^{\infty}$ and $b := \{b_n\}_{n=1}^{\infty}$ be two Cauchy sequences. The **product** of a and b (denoted by $a \times b$ or ab) is defined by

$$ab := \{a_n b_n\}_{n=1}^{\infty}.$$

Definition. (Negation of real numbers) The **negation** of a real number $a := \{a_n\}_{n=1}^{\infty}$ (denoted by $-a$) is defined by

$$-a := \{-1\}_{n=1}^{\infty} \times \{a_n\}_{n=1}^{\infty} = \{-a_n\}_{n=1}^{\infty}.$$

Remark. All definitions above are well defined. Indeed, it suffices to verify the four axioms of equality.

Remark. Analogous to that natural numbers are identified with integers ($n \equiv n \smile 0$) and integers with rational numbers ($a \equiv a//1$), for every $r \in \mathbb{Q}$, we identify r with the real number $\{r\}_{n=1}^{\infty}$, i.e.,

$$r \equiv \{r\}_{n=1}^{\infty} = \{r, r, \cdots\}.$$

Remark. For every real number $a = \{a_n\}_{n=1}^{\infty}$, we can find another Cauchy sequence $\{a'_n\}_{n=1}^{\infty}$ with $a'_n \neq 0$ ($n \in \mathbb{N}$) such that $a = \{a'_n\}_{n=1}^{\infty}$ (WHY).

Indeed, if $a = 0 = \{0\}_{n=1}^{\infty}$, then $a = \{1/n\}_{n=1}^{\infty}$; if $a \neq 0$, refer to the proof of the next proposition.

Problem. Could you give an appropriate definition of the reciprocal of real numbers?

The non-zero real number and its reciprocal

Example. Denote the real number zero by $0 := \{0\}_{n=1}^{\infty}$. Then

$$0 = \{1/n\}_{n=1}^{\infty} = \{1/(2n)\}_{n=1}^{\infty}.$$

Proof. To show the first equality, by definition of real number, it suffices to show that the two Cauchy sequences $\{0\}_{n=1}^{\infty}$ and $\{1/n\}_{n=1}^{\infty}$ are equivalent. To this end, let ϵ be a positive rational number and take $N \in \mathbb{N}$ such that $N > 1/\epsilon$. Then

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} < \epsilon, \quad n \geq N.$$

By analogous argument, we can prove the second equality. □

Proposition. Let $a = \{a_n\}_{n=1}^{\infty}$ be a non-zero real number. Then there is a Cauchy sequence $\{b_n\}_{n=1}^{\infty}$ satisfying the following conditions:

(i) $|b_n| \geq c$ ($n \in \mathbb{N}$) for some positive rational number c.

(ii) $a = \{b_n\}_{n=1}^{\infty}$.

Proof. Note that $a = 0$ if and only if for every rational number $\epsilon > 0$, there is a natural number N such that

$$|a_n| < \epsilon, \quad n \geq N. \quad \text{(WHY)}$$

Thus, $a \neq 0$ if and only if there is a rational number $\epsilon_0 > 0$ such that for every $k \in \mathbb{N}$ there is a natural number $n_k \geq k$ satisfying $|a_{n_k}| \geq \epsilon_0$.

I. Write $b_k := a_{n_k}$ with $k = 1, 2, \cdots$. It is clear that $|b_k| \geq \epsilon_0 := c$ for all $k \in \mathbb{N}$.

II. We show that $\{b_n\}_{n=1}^{\infty}$ is a Cauchy sequence. Indeed, since $\{a_n\}_{n=1}^{\infty}$ is a Cauchy sequence, for every rational number $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that

$$|a_m - a_n| < \epsilon, \quad m, n \geq N.$$

Then for all $k_1, k_2 \geq N$, $n_{k_1} \geq k_1 \geq N$ and $n_{k_2} \geq k_2 \geq N$, and hence,

$$|b_{k_1} - b_{k_2}| = |a_{n_{k_1}} - a_{n_{k_2}}| < \epsilon, \quad k_1, k_2 \geq N.$$

III. We show that $\{b_n\}_{n=1}^{\infty}$ is equal to $\{a_n\}_{n=1}^{\infty}$. Indeed, for all $k \geq N$ with N given in step II, we have $n_k \geq k \geq N$, so that

$$|a_k - b_k| = |a_k - a_{n_k}| < \epsilon.$$

This implies that $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ are equivalent Cauchy sequences. $\square$

Definition. (Non-degenerate Cauchy sequence) A Cauchy sequence $\{a_n\}_{n=1}^{\infty}$ is called **non-degenerate** if there is a rational number $c > 0$ such that $|a_n| \geq c$ for all $n \in \mathbb{N}$.

Remark. The last proposition shows that *every non-zero real number can be expressed by a non-degenerate Cauchy sequence.*

Definition. (Reciprocal of non-zero real numbers) Let $a := \{a_n\}_{n=1}^{\infty}$ be a non-zero real number with $\{a_n\}_{n=1}^{\infty}$, a non-degenerate Cauchy sequence. Then the **reciprocal** of a (denoted by a^{-1}) is defined by

$$a^{-1} := \{a_n^{-1}\}_{n=1}^{\infty}.$$

The system of real numbers $(\mathbb{R}, +, -, \times, \div)$

By using the negation of real numbers, we can construct the operation of substraction for real numbers.

Definition. (Substraction of real numbers) Let $a, b \in \mathbb{R}$. The **substraction** $a - b$ is defined by

$$a - b := a + (-b).$$

Analogously, we can also construct the operation of division for real numbers by using the reciprocal of real numbers.

Definition. (Quotient of real numbers) Let $a, b \in \mathbb{R}$ and $b \neq 0$. The **quotient** a/b is defined by

$$a/b := a \times b^{-1}.$$

Remark. We also use our customary $a \div b$ instead of a/b.

So far, we have constructed the system of real number $(\mathbb{R}, +, -, \times, \div)$. Further on, we can show that the system $(\mathbb{R}, +, -, \times, \div)$ is indeed a so-called **field**.

Theorem. Let $a, b, c \in \mathbb{R}$. Then we have

$$\begin{aligned}
a + b =& b + a, \\
(a + b) + c =& a + (b + c), \\
a + 0 =& 0 + a = a, \\
a + (-a) =& (-a) + a = 0, \\
ab =& ba, \\
(ab)c =& a(bc), \\
a1 =& 1a = a, \\
aa^{-1} =& a^{-1}a = 1 \quad (a \neq 0), \\
a(b + c) =& ab + ac, \\
(b + c)a =& ba + ca.
\end{aligned}$$

Absolute value, order and distance

Definition.(Positive real numbers) A real number a is said to be **positive** if $a = \{a_n\}_{n=1}^{\infty}$, a Cauchy sequence, with $a_n \geq c$ ($n \in \mathbb{N}$) for some rational number $c > 0$; and a is said to be **negative** if $a = -b$ for some positive real number b.

Analogous to that of rational numbers, we have the following trichotomy of real numbers.

Theorem.(Trichotomy of real numbers) Let $a \in \mathbb{R}$. Exactly one of the following three statements is true:

(i) a is zero.

(ii) a is positive.

(iii) a is negative.

Proof. Let $a = \{a_n\}_{n=1}^{\infty}$, a Cauchy sequence of rational numbers a_n. If

(1) for every rational number $r > 0$, there exists an $N \in \mathbb{N}$ such that

$$|a_n - 0| < r$$

for all $n \geq N$, then $a = 0$ due to the equivalence of Cauchy sequences $\{a_n\}_{n=1}^{\infty}$ and $\{0\}_{n=1}^{\infty}$. Otherwise,

(2) there exists a rational number $c > 0$ such that, for every $k \in \mathbb{N}$, there is an a_{n_k} satisfying

$$|a_{n_k}| \geq c.$$

Note that $\{a_{n_k}\}_{k=1}^{\infty}$ is also a Cauchy sequence. Thus, there is a $K \in \mathbb{N}$ such that either

$$a_{n_k} \geq c$$

for all $k \geq K$ or

$$a_{n_k} \leq -c$$

for all $k \geq K$ (WHY). The former implies that $a > 0$ and the latter implies that $a < 0$.

Clearly, we can not have more than one of above statements (i), (ii) and (iii) holding at the same time. This completes the proof. $\square$

Proposition. Let $a \in \mathbb{R}$. The following statements are equivalent.

(i) $a > 0$.

(ii) There is a rational number $c > 0$ and a Cauchy sequence $\{a_n\} = a$ such that $a_n \geq c$ for all $n \in \mathbb{N}$.

(iii) There is a rational number $c > 0$ and a Cauchy sequence $\{a_n\} = a$ such that $a_n \geq c$ $(n \geq N)$ for some $N \in \mathbb{N}$.

(iv) There is a rational number $d > 0$ such that, for every Cauchy sequence $\{a_n\} = a$, there exists an $N' \in \mathbb{N}$ satisfying $a_n \geq d$ for all $n \geq N$.

Proof. The equivalence of (i) and (ii) is a direct consequence of the definition of positivity of real numbers. Others can be proved easily. $\square$

Proposition.(Basic properties of real numbers) Let $a, b \in \mathbb{R}$. The following statements hold.

(i) a is negative **if and only if** $-a$ is positive.

(ii) If a and b are both positive, then so are $a + b$ and ab.

(iii) If a and b are both negative, then so is $a + b$ while ab is positive.

Definition.(Absolute value) Let $a \in \mathbb{R}$. The absolute value of a (denoted by $|a|$) is defined as follows:

(i) $|a| := 0$ if a is zero,

(ii) $|a| := a$ if a is positive,

(iii) $|a| := -a$ if a is negative.

Definition.(Distance of real numbers) Let $a, b \in \mathbb{R}$. The distance $d(a, b)$ is defined by

$$d(a, b) := |a - b|.$$

Definition.(Order of real numbers) Let $a, b \in \mathbb{R}$. We say that a is **greater than** b (denoted by $a > b$) if $a - b$ is positive, and say that a is **less than** b (denoted by $a < b$) if $a - b$ is negative. In addition, we define $a \geq b$ if $a > b$ or $a = b$, and $a \leq b$ if $a < b$ or $a = b$.

Proposition. Let $a, b \in \mathbb{R}$. The following statements are equivalent.

(i) $a > b$.

(ii) There exist a rational number $c > 0$ and two Cauchy sequences $\{a_n\} = a$ and $\{b_n\} = b$ such that

$$a_n - b_n \geq c, \quad n \in \mathbb{N}.$$

(iii) There exist a rational number $c > 0$ and two Cauchy sequences $\{a_n\} = a$ and $\{b_n\} = b$ such that

$$a_n - b_n \geq c \quad (n \geq N)$$

for some $N \in \mathbb{N}$.

(iv) There exists a rational number $d > 0$ such that, for every pair of Cauchy sequences $\{a'_n\} = a$ and $\{b'_n\} = b$, there is an $N' \in \mathbb{N}$ satisfying

$$a_n - b_n \geq d, \quad n \geq N'.$$

Proof. The equivalence of (i) and (ii) is a direct consequence of the definition of positivity of real numbers. Clearly, (ii) implies (iii).

(iii) $\Rightarrow$ (iv) Suppose that the statement (iii) holds. Let $\{a'_n\} = a$ and $\{b'_n\} = b$. By the definition of equivalence of Cauchy sequences, there is a natural number $N' \geq N$ such that $|a'_n - a_n| < c/4$ and $|b'_n - b_n| < c/4$, for all $n \geq N'$. Thus,

$$\begin{aligned} a'_n - b'_n =& a'_n - a_n + a_n - b_n + b_n - b'_n \\ \geq& - c/4 + c - c/4 = c/2 := d > 0, \quad n \geq N'. \end{aligned}$$

(iv) $\Rightarrow$ (ii) Write $a_n := a'_n$ and $b_n := b'_n$ for all $n \geq N$, and write $a_n := d+1$ and $b_n := 1$ for all $n = 1, 2, \cdots, N-1$. $\square$

Theorem.(Basic properties of order) Let $a, b, c \in \mathbb{R}$. The following statements hold.

(i) Exactly one of the following three statements is true:

$$a > b, \quad a = b, \quad a < b.$$

(ii) $a > b$ if and only if $a < b$.

(iii) If $a > b$ and $b > c$, then $a > c$.

(iv) If $a > b$, then $a + c > b + c$.

(v) If $a > b$ and c is positive, then $ac > bc$.

Corollary. Let $a, b \in \mathbb{R}$ be positive. The following statements hold.

(i) a^{-1} is also positive.

(ii) If $a > b$, then $a^{-1} < b^{-1}$.

Proposition. (Basic properties of absolute value) Let $a, b, c \in \mathbb{R}$. The following statements hold.

(i) **(Non-degeneracy)** $0 \leq |a|$. Also, $|a| = 0$ if and only if $a = 0$.

(ii) **(Triangle inequality)** $|a + b| \leq |a| + |b|$.

(iii) $-b \leq a \leq b$ if and only if $|a| \leq b$.

(iv) $|ab| = |a||b|$. In particular, $|-a| = |a|$.

Proposition. (Basic properties of distance) Let $a, b, c \in \mathbb{R}$. The following statements hold.

(i) **(Non-degeneracy)** $0 \leq d(a, b)$. Also, $d(a, b) = 0$ if and only if $a = b$.

(ii) **(Symmetry)** $d(a, b) = d(b, a)$.

(iii) **(Triangle inequality)** $d(a, c) \leq d(a, b) + d(b, c)$.

Remark. Let $r \in \mathbb{Q}$. It is clear that $\{r\}_{n=1}^{\infty}$ is a Cauchy sequence. In the sequel, we will identify the real number $\{r\}_{n=1}^{\infty}$ with the rational number r itself, i,e,, $r \equiv \{r\}_{n=1}^{\infty}$.

Archimedean property

Proposition. Let $a = \{a_n\}_{n=1}^{\infty}$ be a Cauchy sequence of non-negative rational numbers. Then a is a non-negative real number.

Corollary. Let $a = \{a_n\}_{n=1}^{\infty}, b = \{b_n\}_{n=1}^{\infty}$ be two Cauchy sequences. If $a_n \geq b_n$ for all $n \in \mathbb{N}$, then $a \geq b$.

Proposition. Let $a \in \mathbb{R}$ be positive. Then there are two numbers $q \in \mathbb{Q}, N \in \mathbb{Z}$ such that

$$0 < q \leq a \leq N.$$

Corollary.(Archimedean property) Let $a, \epsilon \in \mathbb{R}$ be positive. Then there is an integer $M > 0$ such that $a < M\epsilon$.

Proof. It suffices to show $a/\epsilon < M$ for some integer $M > 0$ (WHY). Since a/ϵ is positive, the existence of such M is a direct consequence of the last proposition. $\square$

Proposition. Let $a, b \in \mathbb{R}$ with $a < b$. Then there is a $q \in \mathbb{Q}$ such that $a < q < b$.

Proof. Since $b > a$, we can take two Cauchy sequences $\{a_n\} = a$ and $\{b_n\} = b$ such that

$$b_n - a_n \geq c \quad (n \in \mathbb{N})$$

for some rational number $c > 0$. Also, by the definition of Cauchy sequences, for the rational number $c/4 > 0$, there is an $N \in \mathbb{N}$ such that

$$|a_n - a_N| < c/4, \quad |b_n - b_N| < c/4,$$

for all $n \geq N$. Let $q := b_N - c/2 \in \mathbb{Q}$. Then, for all $n \geq N$,

$$b_n - q = b_n - b_N + c/2 > -c/4 + c/2 = c/4 > 0$$

and

$$q - a_n = b_N - a_N - c/2 + a_N - a_n > c - c/2 - c/4 > 0.$$

Therefore, $a < q < b$. $\square$

Theory of limits on $\mathbb{R}$

Definition.(Limit) Let $\{a_n\} \subset \mathbb{R}$ and let $a \in \mathbb{R}$. We say that the sequence $\{a_n\}$ **converges to** a (denoted by $\lim_{n\to\infty} a_n = a$) if for every real number $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$|a_n - a| < \epsilon$$

for all $n \geq N$. Also, we say that $\{a_n\}$ is **convergent** if $\{a_n\}$ converges to some $a \in \mathbb{R}$.

Remark. Let $\{a_n\} \subset \mathbb{R}$ and let $a \in \mathbb{R}$. Then $\lim_{n\to\infty} a_n = a$ if and only if one of the following statements holds.

(i) For every real number $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that $|a_n - a| < \epsilon$ for all $n \geq N$.

(ii) For every real number $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that $|a_n - a| \leq \epsilon$ for all $n \geq N$.

(iii) For every rational number $r > 0$, there is an $N \in \mathbb{N}$ such that $|a_n - a| < r$ for all $n \geq N$.

(iv) For every rational number $r > 0$, there is an $N \in \mathbb{N}$ such that $|a_n - a| \leq r$ for all $n \geq N$.

(v) For every natural number $k > 0$, there is an $N \in \mathbb{N}$ such that $|a_n - a| < 1/k$ for all $n \geq N$.

(vi) For every natural number $k > 0$, there is an $N \in \mathbb{N}$ such that $|a_n - a| \leq 1/k$ for all $n \geq N$.

(vii) For every real number $\epsilon = \{\epsilon_k\}_{k=1}^{\infty} > 0$, there is an $N \in \mathbb{N}$ such that $|a_n^k - a^k| < \epsilon_k$ $(k \in \mathbb{N})$ for all $n \geq N$.

(viii) For every real number $\epsilon = \{\epsilon_k\}_{k=1}^{\infty} > 0$, there are $N, K \in \mathbb{N}$ such that $|a_n^k - a^k| < \epsilon_k$ $(k \geq K)$ for all $n \geq N$.

(ix) For every rational number $r > 0$, there are $N, K \in \mathbb{N}$ such that $|a_n^k - a^k| < r$ $(k \geq K)$ for all $n \geq N$.

(x) For every natural number $m > 0$, there are $N, K \in \mathbb{N}$ such that $|a_n^k - a^k| < 1/m$ $(k \geq K)$ for all $n \geq N$.

Theorem.(Cauchy convergence principle) Let $\{a_n\}$ be a sequence of real numbers a_n. The following statements are equivalent.

(i) The sequence $\{a_n\}$ is convergent.

(ii) For every real number $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$|a_m - a_n| < \epsilon$$

for all $m, n \geq N$.

Theorem.(Weierstrass monotone convergence criteria) Let $\{a_n\}$ be a sequence of real numbers a_n.

(i) If $\{a_n\}$ is bounded above and

$$a_1 \leq a_2 \leq \cdots \leq a_n \leq \cdots ,$$

then $\{a_n\}$ is convergent.

(ii) If $\{a_n\}$ is bounded below and

$$a_1 \geq a_2 \geq \cdots \geq a_n \geq \cdots ,$$

then $\{a_n\}$ is convergent.

Theorem.(Cantor criterion of nested intervals) Let $\{I_n\}$ be a sequence of closed intervals $I_n = [a_n, b_n]$. If

$$I_1 \supset I_2 \supset \cdots \supset I_n \supset \cdots ,$$

and $\lim_{n\to\infty}(b_n - a_n) = 0$, then there exists a unique real number $a \in \bigcap_{n=1}^{\infty} I_n$ such that $a = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

Theorem.(Existence of supremum and infimum) Let $A \subset \mathbb{R}$ be non-empty and bounded. Then there exist unique supremum of A and unique infimum of A.

Theorem.(Bolzano-Weierstrass) Let $\{a_n\}$ be a sequence of real numbers a_n. Then there exists a convergent subsequence of $\{a_n\}$ whenever the sequence $\{a_n\}$ is bounded.

Theorem.(Heine-Borel) Let $\mathcal{F}$ be a family of open intervals, and let F be a bounded and closed set satisfying

$$F \subset \bigcup_{O \in \mathcal{F}} O.$$

Then there are finitely many intervals, $O_1, O_2, \cdots, O_m \in \mathcal{F}$, such that

$$F \subset \bigcup_{n=1}^{m} O_n.$$

Summary (Systems of numbers).

- $\mathbb{N}$ is constructed by using so-called "natural" **axioms**.
- $\mathbb{Z}$ is constructed by using $\mathbb{N}$ and $n \equiv n \smile 0$ $(n \in \mathbb{N})$.
- $\mathbb{Q}$ is constructed by using $\mathbb{Z}$ and $a \equiv a//1$ $(a \in \mathbb{Z})$.
- $\mathbb{R}$ is constructed by using $\mathbb{Q}$ and $r \equiv \{r\}_{n=1}^{\infty}$ $(r \in \mathbb{Q})$.
- $\mathbb{R}$ is **complete**.
- $\mathbb{C}$ is constructed by using $\mathbb{R}$ and $a \equiv (a, 0)$ $(a \in \mathbb{R})$.
- $\mathbb{C}$ is **complete**.
- Theory of series $\sum_{n=1}^{\infty} a_n$, where $\{a_n\} \subset \mathbb{R}$.
- Base-p expansions.

Chapter 1: Theory of sets

§1. Basics of sets

Operations of sets

- **Union** $A \cup B := \{x : x \in A \text{ or } x \in B\}$
  $\cup_{\alpha\in\Lambda} A_\alpha = \{x : x \in A_\alpha \text{ for some } \alpha \in \Lambda\}$
- **Intersection** $A \cap B := \{x : x \in A \text{ and } x \in B\}$
  $\cap_{\alpha\in\Lambda} A_\alpha = \{x : x \in A_\alpha \text{ for all } \alpha \in \Lambda\}$
- **Difference** $A \backslash B := \{x : x \in A \text{ and } x \notin B\}$
- **Complement** $A^c := X \setminus A$
- **Symmetric difference** $A \triangle B := (A \setminus B) \cup (B \setminus A)$

Properties.

(1) $A \cup A = A$, $A \cap A = A$
(2) $A \triangle A = \emptyset, A \triangle X = A^c, A \triangle B = (A \cup B) \setminus (A \cap B)$
(3) $A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$, $A \setminus \emptyset = A, A \triangle \emptyset = A$
(4) $A \cup B = B \cup A$, $A \cap B = B \cap A, A \triangle B = B \triangle A$
(5) $A \cup (B \cup C) = (A \cup B) \cup C$
$A \cap (B \cap C) = (A \cap B) \cap C$
$A \triangle (B \triangle C) = (A \triangle B) \triangle C$
(6) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
(7) $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$
(8) $(A \backslash B) \backslash C = A \backslash (B \cup C), A \backslash (B \backslash C) = (A \backslash B) \cup (A \cap C)$
(9) $A \cup B = (A \Delta B) \cup (A \cap B) = A \cup (B \setminus A)$
(10) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$
(11) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$

Proposition. (**de Morgan**) Let S be a set and let $\{A_\alpha : \alpha \in \Lambda\}$ be a family of sets. Then

$$S \setminus \bigcup_{\alpha\in\Lambda} A_\alpha = \bigcap_{\alpha\in\Lambda} (S \setminus A_\alpha), \quad S \setminus \bigcap_{\alpha\in\Lambda} A_\alpha = \bigcup_{\alpha\in\Lambda} (S \setminus A_\alpha).$$

Rings and algebras

A **ring**, $\mathcal{R}$, over a set X, is a class of subsets of X such that $A \cup B, A \setminus B \in \mathcal{R}$ for all $A, B \in \mathcal{R}$. An **algebra** $\mathcal{F}$ over X is a ring over X with $X \in \mathcal{F}$; an algebra is also called a **field**.

Remarks. Let $\mathcal{R}$ be a ring over a set X.

(i) It is clear that $\emptyset \in \mathcal{R}$. In particular, $\mathcal{R} = \{\emptyset\}$ whenever $X = \emptyset$.

(ii) $\mathcal{R}$ is closed under the operation intersection.

Proposition. Let $\mathcal{R}$ be a class of subsets of a set X. Then $\mathcal{R}$ is a ring over X if and only if $\mathcal{R}$ is closed under operations of finitely many unions and differences.

Proposition. Let $\mathcal{F}$ be a class of subsets of a set X. The following statements are equivalent.

(i) $\mathcal{F}$ is an algebra over X.

(ii) $\mathcal{F}$ is closed under operations of union, difference and complement.

(iii) $\mathcal{F}$ is closed under operations of finitely many unions, differences and complements.

Examples.

(1) Let $\mathcal{P}(X) := \{A : A \subset X\}$, then $\mathcal{P}(X)$ is the largest ring/algebra over X. $\{\emptyset\}$ and $\{\emptyset, X\}$ are the smallest ring and algebra over X, respectively.

(2) Let $X = \{1, 2\}$. Then

$$\mathcal{R}_1 = \big\{\emptyset, \{1\}\big\} \text{ and } \mathcal{R}_2 = \big\{\emptyset, \{2\}\big\}$$

are both rings on X.

(3) Let $\mathcal{F}$ be the class of all finitely many unions of semi-closed "intervals" of the form $(a, b]$, then $\mathcal{F}$ is a *ring* over $\mathbb{R}$.

Note that $\mathcal{R}_1 \cup \mathcal{R}_2$ is not a ring with $\mathcal{R}_1, \mathcal{R}_2$ given in above Example (2). However, $\mathcal{R}_1 \cap \mathcal{R}_2$ is a ring. In general, we have the following

Proposition. Let $\mathcal{M}$ be a family of rings over X. Then

$$\mathcal{R} := \bigcap_{\mathcal{R}' \in \mathcal{M}} \mathcal{R}'$$

is a ring over X as well.

Proposition. Let $\mathcal{M}$ be a family of algebras over X. Then

$$\mathcal{F} := \bigcap_{\mathcal{F}' \in \mathcal{M}} \mathcal{F}'$$

is an algebra over X as well.

Proposition. Let $A \in \mathcal{P}(X)$. Then there exists the **smallest** ring (or algebra) $\mathcal{R}$ (or $\mathcal{F}$) over X such that $A \in \mathcal{R}$ (or $\mathcal{A} \subset \mathcal{F}$), that is to say, for every ring $\mathcal{R}'$ (or every algebra $\mathcal{F}'$) over X satisfying $A \in \mathcal{R}'$ (or $A \in \mathcal{F}'$) we have $\mathcal{R} \subset \mathcal{R}'$ (or $\mathcal{F} \subset \mathcal{F}'$).

Proof. Indeed, $\mathcal{R} = \{\emptyset, A\}$ and $\mathcal{F} = \{\emptyset, A, A^c, X\}$. $\square$

Theorem. Let $\mathcal{A} \subset \mathcal{P}(X)$. Then there exists the **smallest** ring (or algebra) $\mathcal{R}$ (or $\mathcal{F}$) over X such that $\mathcal{A} \subset \mathcal{R}$ (or $\mathcal{A} \subset \mathcal{F}$), that is to say, for every ring $\mathcal{R}'$ (or every algebra $\mathcal{F}'$) over X satisfying $\mathcal{A} \subset \mathcal{R}'$ (or $\mathcal{A} \subset \mathcal{F}'$) we have $\mathcal{R} \subset \mathcal{R}'$ (or $\mathcal{F} \subset \mathcal{F}'$).

Proof. Note that the set $\mathcal{F}$ defined by

$$\mathcal{F} := \bigcap \big\{\mathcal{F}' : \mathcal{A} \subset \mathcal{F}' \subset \mathcal{P}(X) \text{ and } \mathcal{F}' \text{ is an algebra}\big\}$$

is an algebra over X. Clearly, $\mathcal{F}$ is the smallest algebra containing A as an element. Also,

$$\mathcal{R} := \bigcap \{\mathcal{R}' : \mathcal{A} \subset \mathcal{R}' \subset \mathcal{P}(X) \text{ and } \mathcal{R}' \text{ is a ring}\}$$

is the ring desired. $\square$

Remark. $\mathcal{R}$ (or $\mathcal{F}$) given in above theorem is called the **ring (or algebra) generated by** A, and we denote it by $\mathscr{R}(A)$ (or $\mathscr{F}(A)$). In particular, $\mathscr{R}(\emptyset) = \{\emptyset\}$ and

$$\mathscr{R}(X) = \mathscr{F}(\emptyset) = \{\emptyset, X\}.$$

Limits of sets

Monotonic sequences of sets.

(1) $\{A_n\}$ is **increasing**: $A_n \subset A_{n+1}$ for all n.

(2) $\{A_n\}$ is **decreasing**: $A_n \supset A_{n+1}$ for all n.

(3) $\{A_n\}$ is **monotone** if it is either increasing or decreasing.

Remark. Let $\{A_n\}$ be decreasing. Then

$$A_1 = (A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup \cdots \cup (A_n \setminus A_{n+1}) \cup \cdots \cup \left(\bigcap_{n=1}^{\infty} A_n \right);$$

moreover, all terms are pairwise disjoint.

Limits of sets.

(1) **limit superior**:

$$\overline{\lim_{n\to\infty}} A_n := \limsup_{n\to\infty} A_n := \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$$

(2) **limit inferior**:

$$\underline{\lim_{n\to\infty}} A_n := \liminf_{n\to\infty} A_n := \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n$$

(3) **limit**: We say that the sequence $\{A_n\}$ of sets A_n is **convergent** if

$$\limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n,$$

and define the **limit** of $\{A_n\}$ by

$$\lim_{n\to\infty} A_n := \limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n.$$

Properties.

(1) $\overline{\lim\limits_{n\to\infty}} A_n = \{x : x \in A_n \text{ for infinitely many } n\}$.
(2) $\underline{\lim\limits_{n\to\infty}} A_n = \{x : x \in A_n \text{ for all but finitely many } n\}$.
(3) $\bigcap\limits_{n=1}^{\infty} A_n \subset \underline{\lim\limits_{n\to\infty}} A_n \subset \overline{\lim\limits_{n\to\infty}} A_n \subset \bigcup\limits_{n=1}^{\infty} A_n$.

Proposition.

(1) If $\{A_n\}$ is increasing, then

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n.$$

(2) If $\{A_n\}$ is decreasing, then

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n.$$

Examples. Let $n \in \mathbb{N}^+$.

(1) Let $A_n = [n, \infty)$. Then $\lim_{n\to\infty} A_n = \emptyset$.

(2) Let $A_n = (-\frac{1}{n}, \frac{1}{n})$. Then $\lim_{n\to\infty} A_n = \{0\}$.

(3) Let $A_n = (-1 + \frac{1}{n}, 1 - \frac{1}{n}), n \geq 2$. Then

$$\lim_{n\to\infty} A_n = (-1, 1).$$

(4) Let $A_{2n} = [0, 2 - \frac{1}{2n+1}], A_{2n+1} = [0, 1 + \frac{1}{2n}]$. Then

$$\overline{\lim_{n\to\infty}} A_n = [0, 2), \quad \underline{\lim_{n\to\infty}} A_n = [0, 1].$$

(5) Let $f_n, f : \mathbb{R} \to \mathbb{R}$. Then

$$\{t : f_n(t) \to f(t)\} = \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \left\{t : |f_n(t) - f(t)| < \frac{1}{k}\right\},$$

$$\{t : f_n(t) \not\to f(t)\} = \bigcup_{k=1}^{\infty} \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \left\{t : |f_n(t) - f(t)| \geq \frac{1}{k}\right\}.$$

Functions and sets.

Let $f : E \to \mathbb{R}$ (or $\mathbb{C}$). Write

$$\begin{aligned} E[f \geq c] :=& \{x \in E : f(x) \geq c\}, \\ E[f \leq c] :=& \{x \in E : f(x) \leq c\}, \\ E[f < c] :=& \{x \in E : f(x) < c\}, \\ E[f > c] :=& \{x \in E : f(x) > c\}, \\ E[f = c] :=& \{x \in E : f(x) = c\}. \end{aligned}$$

It is clear that

(1) $E[f \geq c] \cup E[f < c] = E$.
(2) $E[f \geq c] \cap E[f < c] = \emptyset$.
(3) $E[f > c] \cap E[f \leq d] = E[c < f \leq d]$.
(4) $E[f \leq c] = \bigcap_{n=1}^{\infty} E[f < c + \frac{1}{n}]$.

Proposition. Let $f_n \to f$ as $n \to \infty$ (i.e., pointwise). Then

$$\begin{aligned} E[f \leq c] &= \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n \geq N}^{\infty} E[f_n \leq c + 1/k], \\ &= \bigcap_{k=1}^{\infty} \varliminf_{n \to \infty} E[f_n \leq c + 1/k]. \end{aligned}$$

If further $f_1 \leq f_2 \leq \cdots \leq f_n \leq f_{n+1} \leq \cdots$, then

$$E[f \leq c] = \bigcap_{n=1}^{\infty} E[f_n \leq c] = \lim_{n \to \infty} E[f_n \leq c].$$

Proof. It suffices to prove the first equality. Let

$$x \in \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n \geq N}^{\infty} E[f_n \leq c + 1/k].$$

Then, for every $k \in \mathbb{N}$, there is an $N \in \mathbb{N}$ such that

$$x \in E[f_n \leq c + 1/k]$$

for all $n \geq N$, i.e.,

$$f_n(x) \leq c + 1/k, \quad n \geq N.$$

By letting $n \to \infty$ in both sides of above inequality, we obtain

$$f(x) \leq c + 1/k, \quad k \in \mathbb{N}.$$

Again, letting $k \to \infty$ yields $f(x) \leq c$. Thus, $x \in E[f \leq c]$.

Conversely, let $x \in E[f \leq c]$. Then

$$a := \lim_{n\to\infty} f_n(x) = f(x) \leq c.$$

Thus, for every $k \in \mathbb{N}$, there is an $N \in \mathbb{N}$ such that

$$|f_n(x) - a| \leq 1/k, \quad n \geq N,$$

and hence,

$$f_n(x) - c \leq f_n(x) - a \leq 1/k, \quad n \geq N.$$

This implies that

$$x \in \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n\geq N}^{\infty} E[f_n \leq c + 1/k].$$

$\square$

Characteristic functions of sets.

Let X be a nonempty set and let $A \subset X$. We call the function χ_A (defined on X)

$$\chi_A(x) := \begin{cases} 1, & x \in A; \\ 0, & x \notin A, \end{cases}$$

the characteristic function of A.

We list some basic properties as follows, which can be proved directly by the definition of characteristic functions and the de Morgan law of sets.

(1) $\chi_A \equiv 1 \Leftrightarrow A = X; \quad \chi_A \equiv 0 \Leftrightarrow A = \emptyset$.

(2) $\chi_A \leq \chi_B \Leftrightarrow A \subset B; \quad \chi_A = \chi_B \Leftrightarrow A = B$.

(3) Let $U := \bigcup_{\alpha\in\Lambda} A_\alpha$ and $M := \bigcap_{\alpha\in\Lambda} A_\alpha$. Then

$$\chi_U(x) = \max_{\alpha\in\Lambda} \chi_{A_\alpha}(x),$$
$$\chi_M(x) = \min_{\alpha\in\Lambda} \chi_{A_\alpha}(x).$$

(4) Let $\{A_n\} \subset X$. Then, for every $x \in X$,

$$\chi_{\overline{\lim}_{n\to\infty} A_n}(x) = \overline{\lim_{n\to\infty}} \chi_{A_n}(x),$$
$$\chi_{\underline{\lim}_{n\to\infty} A_n}(x) = \underline{\lim_{n\to\infty}} \chi_{A_n}(x).$$

(5) Let $\{A_n\} \subset X$. Then $\{A_n\}$ converges if and only if $\{\chi_{A_n}(x)\}$ converges for every $x \in X$. Moreover,

$$\chi_{\lim_{n\to\infty} A_n}(x) = \lim_{n\to\infty} \chi_{A_n}(x), \quad x \in X.$$

Proof. (2) Suppose that $\chi_A \leq \chi_B$, i.e.,

$$\chi_A(x) \leq \chi_B(x), \quad x \in X.$$

Let $x \in A$. Then $\chi_B(x) \geq \chi_A(x) = 1$, so that $\chi_B(x) = 1$. This implies that $x \in B$. Thus, $A \subset B$.

Conversely, suppose that $A \subset B$.

(i) If $x \in A$, then $x \in B$, so that $\chi_A(x) = \chi_B(x) = 1$.

(ii) If $x \notin A$, then $\chi_A(x) = 0 \leq \chi_B(x)$.
In a word, we have $\chi_A \leq \chi_B$.

Analogously, (3) and (4) can be proved by the de Morgan law of sets. (5) is a direct consequence of (4). $\square$

§2. Mappings and Cardinalities of sets

Mappings.

Some notations.

Let A, B be nonempty sets. A **mapping** φ from A to B is a rule that assigns to each element $x \in A$ a unique element $\varphi(x) \in B$. We call $\varphi(x)$ the **image** of x under the mapping φ.

To indicate that φ is a mapping from A to B, we often write $\varphi : A \to B$. The set A is called the **domain** of φ (denoted by $D(\varphi)$). The set $\varphi(A) := \{\varphi(x) : x \in A\}$ is called the **range** of φ (also, denoted by $R(\varphi)$). Clearly, $\varphi(A) \subset B$.

For $y \in B$ fixed, the set $\{x \in A : \varphi(x) = y\}$ is called the **inverse image** of y under the mapping φ (denoted by $\varphi^{-1}(y)$). In addition, for $E \subset B$ fixed, the set $\{x \in A : \varphi(x) \in E\}$ is called the **inverse image** of E under the mapping φ (denoted by $\varphi^{-1}(E)$).

The mapping $\varphi : A \to B$ is said to be

(1) **one-to-one** (or **injective**) if $\varphi(x_1) = \varphi(x_2)$ implies that $x_1 = x_2$.
(2) **onto** (or **surjective**) if $\varphi(A) = B$.
(3) **1-1 correspondence** (or **bijective**) if φ is one-to-one and onto.

Remark. Clearly, an injective φ is bijective from $D(\varphi)$ to $R(\varphi)$.

Extensions of mappings. Let $\varphi : D(\varphi) \to B$ and $\psi : D(\psi) \to B$. If $D(\varphi) \subset D(\psi)$ and $\psi(x) = \varphi(x)$ for all

$x \in D(\varphi)$, then we call ψ is an **extension** of φ to $D(\psi)$ (denoted by $\varphi \subset \psi$). If this is the case, we call φ the **restriction** of ψ on $D(\varphi)$ (denoted by $\varphi = \psi|_{D(\varphi)}$).

Compositions of mappings. Let $\varphi_1 : A \to B$ and $\varphi_2 : B \to C$. The **composition** of φ_2 with φ_1, denoted by $\varphi_2 \circ \varphi_1$, is the mapping $\varphi_2 \circ \varphi_1 : A \to C$ defined by

$$(\varphi_2 \circ \varphi_1)(x) = \varphi_2(\varphi_1(x)), \quad x \in A.$$

Inverse of a mapping. Let $\varphi : A \to B$ be injective. For $y \in R(\varphi)$, let $\varphi^{-1}(y)$ be the unique $x \in A$ such that $\varphi(x) = y$. The mapping $\varphi^{-1} : R(\varphi) \to A$ so defined is called the **inverse** of the mapping φ. In this sense, an injective is also said to be an **invertible mapping**.

Identity mapping. Let A be a nonempty set. The **identity mapping** on A is the mapping $\varphi : A \to A$ defined by

$$\varphi(x) = x, \quad x \in A.$$

Equivalence of sets

Definition. Two non-empty sets are said to be **equivalent** if there is a 1-1 correspondence from one to the other. We write $A \sim B$ if sets A and B are equivalent. In addition, we define $\emptyset \sim \emptyset$.

Remark. Let A, B and C be three sets. It is clear that

(i) $A \sim A$ (reflexive)

(ii) $A \sim B$ implies that $B \sim A$ (symmetric)

(iii) $A \sim B$ and $B \sim C$ implies that $A \sim C$ (transitive)

Remark. Let $\varphi : A \to B$ be injective. Then $A \sim R(\varphi)$.

Examples. Let $a, b \in \mathbb{R}$. Then

$$[0,1] \sim [a,b] \sim [a,b) \sim (a,b] \sim (a,b).$$

Theorem (F. Bernstein, 1898). Let A and B be two sets. If A is equivalent to a subset of B while B is equivalent to a subset of A, then A and B are equivalent.

Outline of the proof. Suppose that $A \sim B_1 \subset B$ and $B \sim A_1 \subset A$.

Step I. Since the relation $\sim$ has translation, it suffices to show $A \sim A_1$.

Step II. Consider the following disjoint decompositions

$$A = C_1 \cup C_2 \cup \cdots \cup C_n \cup \cdots,$$
$$A_1 = C'_1 \cup C'_2 \cup \cdots \cup C'_n \cup \cdots.$$

If $C_n \sim C'_n$ for every $n \in \mathbb{N}$, then $A \sim A_1$.

Step III. Construct such decompositions.

Proof. Suppose that $A \sim B_1 \subset B$ and $B \sim A_1 \subset A$ via 1-1 correspondences φ_1 and φ_2, respectively. Write $A_0 := A, \varphi := \varphi_2 \circ \varphi_1$ and denote $A_{n+2} := \varphi(A_n)$ for $n \in \mathbb{N}$. Then we obtain a decreasing series of sets

$$A \supset A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots.$$

Also,

$$A \sim A_2 \sim A_4 \sim \cdots; \quad A_1 \sim A_3 \sim A_5 \sim \cdots$$

due to the same mapping φ. Notice that

$$A = (A \setminus A_1) \cup (A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup (A_3 \setminus A_4) \cup \cdots \cup D,$$
$$A_1 = (A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup (A_3 \setminus A_4) \cup (A_4 \setminus A_5) \cup \cdots \cup D,$$

where $D = A_1 \cap A_2 \cap A_3 \cap \cdots$; and notice that

$$\begin{gathered} A \setminus A_1 \sim A_2 \setminus A_3 \\ A_2 \setminus A_3 \sim A_4 \setminus A_5 \\ \cdots \\ A_{2n} \setminus A_{2n+1} \sim A_{2n+2} \setminus A_{2n+3} \\ \cdots \end{gathered}$$

by virtue of the same mapping φ. This implies that $A \sim A_1$, so that $A \sim B$ due to $A_1 \sim B$. $\square$

Corollary. Let $C \subset B \subset A$. If $C \sim A$ then $B \sim A$.

Cardinalities of sets.

Definition. Let A and B are two sets. We define the expressions

$$\overline{\overline{A}} \leq \overline{\overline{B}} \text{ (or } \overline{\overline{B}} \geq \overline{\overline{A}}\text{) and } \overline{\overline{A}} = \overline{\overline{B}}$$

to mean that A is equivalent to a subset of B and A is equivalent to B, respectively. In addition, we call $\overline{\overline{A}}$ the **cardinality** of A.

Definition. We also write

$$\overline{\overline{A}} < \overline{\overline{B}} \text{ (or } \overline{\overline{B}} > \overline{\overline{A}}\text{)}$$

to mean that $\overline{\overline{A}} \leq \overline{\overline{B}}$ but $\overline{\overline{A}} \neq \overline{\overline{B}}$.

Bernstein theorem. Let A and B be two sets. If $\overline{\overline{A}} \leq \overline{\overline{B}}$ and $\overline{\overline{B}} \leq \overline{\overline{A}}$ then $\overline{\overline{A}} = \overline{\overline{B}}$.

Remark. By **Bernstein theorem** and **Zermelo Axiom of Choice** (introduced in the next section), for two sets A and B, exactly one of the following three statements holds:

$$\overline{\overline{A}} < \overline{\overline{B}}, \quad \overline{\overline{A}} = \overline{\overline{B}}, \quad \overline{\overline{A}} > \overline{\overline{B}}.$$

Examples.

1. $\{a_1, \cdots, a_n\} \sim \{1, \cdots, n\}$.
2. $\{0, 2, 4, \cdots, 2n, \cdots\} \sim \{0, 1, 2, \cdots, n, \cdots\}$.
3. $\mathbb{N} \sim \mathbb{Z}$.
4. $[0, 1] \sim [a, b]$, where $a, b \in \mathbb{R}$ and $a < b$.

Finite and infinite sets:

Definition. A set A is called a **finite set** if $A \sim \{1, \cdots, n\}$ for some $n \in \mathbb{N}$. A set is called an **infinite set** if it is not finite.

Theorem. A set is infinite if and only if it is equivalent to some of its proper subsets, or equivalently, a set is finite if and only if it is not equivalent to any of its proper subsets.

Proof. Let A be an infinite set. We can take

$$\{a_1, a_2, \cdots, a_n, \cdots\} \subset A.$$

Now consider the proper subset $A \setminus \{a_1\}$ and define a mapping $\varphi : A \to A \setminus \{a_1\}$ by

(i) $\varphi(x) := x$ for $x \in A \setminus \{a_1, a_2, \cdots, a_n, \cdots\}$,

(ii) $\varphi(a_k) := a_{k+1}$ for $k = 1, 2, \cdots$.

Clearly, φ is a bijective. Thus, $A \sim A \setminus \{a_1\}$.

Conversely, let B be a proper subset of A such that $B \sim A$. It is clear that $B \neq \emptyset$ (WHY). Assume that A is finite. Then, by definition of finite sets, there exists an $N \in \mathbb{N}$ such that $\{1, 2, \cdots, N\} \sim A$. Thus, there is a bijective $\varphi_1 : \{1, 2, \cdots, N\} \to A$, and hence,

$$A = \{\varphi_1(1), \varphi_1(2), \cdots, \varphi_1(N)\}.$$

Note that

$$B = \{\varphi(n_1), \varphi(n_2), \cdots, \varphi(n_k)\}, \quad 1 \leq k < N.$$

Thus, any mapping from A to B is not an injective (otherwise, $k = N$). This contracts to that $A \sim B$. Thus, A is infinite. $\square$

Countable sets.

Definition. A set A is said to be **countable** if $A \sim \mathbb{N}$.

Remark. For convenience, we write $\overline{\overline{N}} = \aleph_0$. Thus, $\overline{\overline{A}} = \aleph_0$ for every countable set A.

Properties.

1. Every infinite set has a countable subset.

2. Any subset of a countable set is finite or countable.

3. If A is countable and B is finite or countable, then $A \cup B$ is countable.

4. If A_n is countable for every $n \in \mathbb{N}$, then $\bigcup_{n=1}^{\infty} A_n$ is also countable.

5. If A is an infinite set and B is finite or countable, then $A \cup B \sim A$.

Proof. 1. Let A be an infinite set. Clearly, we can take $a_1 \in A$. Since $A \setminus \{a_1\}$ is also infinite (WHY), we can take $a_2 \in A \setminus \{a_1\}$. Note that, if we take $a_1, a_2, \cdots, a_k \in A$, then we can also take $a_{k+1} \in (A \setminus \{a_1, a_2, \cdots, a_k\})$ due to the infinity of A. Thus, by induction, we obtain a countable subset $\{a_1, a_2, \cdots, a_n, \cdots\}$ of A.

2. Let A be a countable set, and let $B \subset A$. It suffices to show that B is countable whenever B is not finite. Suppose that B is not finite. Note that

$$\aleph_0 \leq \overline{\overline{B}} \leq \overline{\overline{A}} = \aleph_0,$$

where the first inequality follows from 1, immediately. This implies that $\overline{\overline{B}} = \aleph_0$ by Bernstein theorem.

3. Let A be countable, and let B be finite or countable. Write

$$A = \{a_1, a_2, \cdots, a_n, \cdots\}.$$

(i) If B is finite, then $B = \{b_1, b_2, \cdots, b_n\}$, so that the mapping φ defined by

$$\begin{aligned}
&\varphi(b_k) := a_k, \quad k = 1, 2, \cdots, n; \\
&\varphi(a_m) := a_{m+n}, \quad m \in \mathbb{N},
\end{aligned}$$

is a bijective from $A \cup B$ to A. Thus, $A \sim B$.

(ii) If B is countable, then $B = \{b_1, b_2, \cdots, b_n, \cdots\}$. Note that the mapping φ defined by

$$\begin{aligned}
&\varphi(a_n) := a_{2n}, \\
&\varphi(b_n) := a_{2n-1}, \quad n \in \mathbb{N},
\end{aligned}$$

is a bijective from $A \cup B$ to A. That is to say, $A \cup B$ is arrayed as

$$b_1, a_1, b_2, a_2, \cdots, b_n, a_n, \cdots.$$

Thus, $A \cup B$ is also countable.

4. Let A_n be countable for every $n \in \mathbb{N}$. Suppose that $A_n = \{a_n^1, a_n^2, \cdots, a_n^m, \cdots\}$. Then $\bigcup_{n=1}^{\infty} A_n$ can be arrayed as

$$\begin{array}{l}
a_1^1, a_1^2, a_1^3, \cdots, a_1^m, \cdots \\
a_2^1, a_2^2, a_2^3, \cdots, a_2^m, \cdots \\
a_3^1, a_3^2, a_3^3, \cdots, a_3^m, \cdots \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
a_n^1, a_n^2, a_n^3, \cdots, a_n^m, \cdots \\
\cdots\cdots\cdots\cdots\cdots\cdots
\end{array}$$

Thus, $\bigcup_{n=1}^{\infty} A_n$ is countable as well.

Finally, the statement 5 can be proved by arguments analogous to that given in the last theorem. $\square$

Examples of countable sets.

1. $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$.
2. $\{(x, y) \in \mathbb{R}^2 : x, y \in \mathbb{Q}\}$.

Uncountable sets.

Definition. An infinite set A is called **uncountable** if $A \not\sim \mathbb{N}$.

Remark. Clearly, $\overline{\overline{A}} > \aleph_0$ for every uncountable set A.

Example. (**Cantor**, 1874) $\overline{\overline{[0,1]}} > \aleph_0$.

Proof. It is clear that $[0, 1]$ is infinite. Assume that $\overline{\overline{[0,1]}} = \aleph_0$. Then we can array it as

$$[0, 1] = \{a_1, a_2, \cdots, a_n, \cdots\}.$$

Trisect the interval $[0, 1]$ and take one, $[b_1, b_1']$, of the three subintervals such that $a_1 \notin [b_1, b_1']$; Trisect the interval $[b_1, b_1']$ and take one, $[b_2, b_2']$, of the three subintervals such that $a_2 \notin [b_2, b_2']$; $\cdots$; Trisect the interval $[b_k, b_k']$ and take one, $[b_{k+1}, b_{k+1}']$, of the three subintervals such that $a_{k+1} \notin [b_{k+1}, b_{k+1}']$. Thus, we obtain a decreasing sequence $\{[b_n, b_n']\}$ with $a_n \notin [b_n, b_n']$ for every $n \in \mathbb{N}$, and hence

$$\{a_1, a_2, \cdots, a_n, \cdots\} \cap \left(\bigcap_{n=1}^{\infty} [b_n, b_n']\right) \neq \emptyset.$$

On the other hand, by the completeness of real numbers, there is a real number $a \in [0, 1]$ such that

$$\bigcap_{n=1}^{\infty} [b_n, b_n'] = \{a\}.$$

Thus, $a \neq a_n$ for all $n \in \mathbb{N}$, so that

$$[0, 1] \neq \{a_1, a_2, \cdots, a_n, \cdots\}.$$

This yields a contradiction. Thus, $\overline{\overline{[0,1]}} > \aleph_0$. $\square$

Cardinality of the continuum. A set A is said to have the cardinality of the continuum if $A \sim [0,1]$. And we denote it by $\bar{\bar{A}} = \aleph$.

Example. Let $a, b \in \mathbb{R}$ and $a < b$. Then

$$\overline{\overline{[a,b]}} = \overline{\overline{[a,b)}} = \overline{\overline{(a,b]}} = \overline{\overline{(a,b)}} = \overline{\overline{\mathbb{R}}} = \aleph.$$

Proposition. Let $\overline{\overline{A_n}} = \aleph$ for every $n \in \mathbb{N}$. Then

$$\bigcup_{n=1}^{\infty} A_n = \aleph.$$

Proof. It is clear that

$$\overline{\overline{\bigcup_{n=1}^{\infty} A_n}} \geq \overline{\overline{A_1}} = \aleph.$$

On the other hand, write

$$A_1' := A_1, A_2' := A_2 \setminus A_1, \cdots, A_n' := A_n \setminus \bigcup_{k<n} A_k, \cdots.$$

Then $\{A_n'\}$ is a disjoint decomposition of $\bigcup_{n=1}^{\infty} A_n$. Note that we can map A_1' to $[1,2)$ via a bijective and A_k' to $[k, k+1)$ via an injective for every $k \geq 2$. Thus, there is an injective from $\bigcup_{n=1}^{\infty} A_n'$ to $[1,\infty)$, and hence,

$$\overline{\overline{\bigcup_{n=1}^{\infty} A_n}} = \overline{\overline{\bigcup_{n=1}^{\infty} A_n'}} \leq \overline{\overline{[1,\infty)}} = \aleph.$$

$\square$

Examples.

1. $\overline{\overline{\mathbb{R}}} = \aleph$.
2. $\overline{\overline{\mathbb{R} - \mathbb{Q}}} = \aleph$.
3. $\overline{\overline{\mathbb{R}^n}} = \overline{\overline{\mathbb{R}^\infty}} = \aleph$.
4. $\overline{\overline{\mathbb{Q}^n}} = \aleph_0$, $\overline{\overline{\mathbb{Q}^\infty}} = \aleph$.
5. $\overline{\overline{\{\{a_n\} : a_n = 0 \text{ or } 1\}}} = \aleph$.

Theorem. (**Cantor**, 1891) Let A be a set. Then

$$\overline{\overline{A}} < \overline{\overline{\mathcal{P}(A)}}.$$

Proof. Suppose to the contrary that there is a bijection ϕ from A to $\mathcal{P}(A)$. Note that $\phi(x) \in \mathcal{P}(A)$. We can define

$$B = \{x \in A : x \notin \phi(x)\},$$

then $B \subset A$, i.e., $B \in \mathcal{P}(A)$. Since ϕ is surjective, there is some $x_0 \in A$ such that $\phi(x_0) = B$. However, either $x_0 \in B$ or $x_0 \notin B$ will give contradiction. $\square$

Remark. Note that, if $\overline{\overline{A}} = n$, then $\overline{\overline{\mathcal{P}(A)}} = 2^n$. Thus, we write $\overline{\overline{\mathcal{P}(A)}} := 2^{\overline{\overline{A}}}$ for short.

Problem. By Cantor theorem, $\overline{\overline{\mathbb{N}}} < 2^{\overline{\overline{\mathbb{N}}}}$, i.e., $\aleph_0 < 2^{\aleph_0}$. Also, we have proved that $\aleph_0 < \aleph$. What about the cardinalities $2^{\aleph_0}$ and $\aleph$?

Proposition. $\aleph_0 < \mathcal{P}(\mathbb{N}) = \aleph$.

Proof. It suffices to show $\overline{\overline{\mathcal{P}(\mathbb{N})}} = \aleph$. We decompose the proof into the following three steps.

(i) For the class

$$\mathcal{B} := \{B \subset \mathbb{N} : B \text{ is finite}\},$$

we have $\overline{\overline{\mathcal{B}}} = \aleph_0$. Indeed, note that

$$\mathcal{B} = \bigcup_{n\in\mathbb{N}} \mathcal{B}_n,$$

where $\mathcal{B}_n$ denotes the class of subsets B with $\overline{\overline{B}} = n$. By the method of induction, it is easy to see that the class $\mathcal{B}_n$ is countable for every $n \in \mathbb{N}$, so that $\mathcal{B}$ is countable.

(ii) For the class

$$\mathcal{C} := \{C \subset \mathbb{N} : C \text{ is infinite}\},$$

we have $\overline{\overline{\mathcal{C}}} = \aleph$. Indeed, for every $x \in \mathcal{C}$ and define the **binary decimal** $\varphi(x)$ as

$$\varphi(x) := 0.a_1a_2\cdots a_n\cdots,$$

where $a_n = 1$ if $n \in C$, and $a_n = 0$ if $n \notin C$. Then φ is a bijective from $\mathcal{C}$ to the set of all **binary infinite decimals** in $(0, 1]$, and hence, $\overline{\overline{\mathcal{C}}} = \aleph$.

(iii) $\overline{\overline{\mathcal{B} \cup \mathcal{C}}} = \aleph$. Indeed,

$$\mathcal{B} \cup \mathcal{C} = (\mathcal{C} \setminus \mathcal{B}) \cup \mathcal{B}.$$

Since $\mathcal{B}$ is countable, there is an injective from $\mathcal{B}$ to $[0, 1)$. Also, since $\overline{\overline{\mathcal{C} \setminus \mathcal{B}}} = \aleph$ (WHY), there is a bijective from $\mathcal{C}\setminus\mathcal{B}$ to $[1, 2)$. Thus, there is an injective from $(\mathcal{C}\setminus\mathcal{B})\cup\mathcal{B}$ to $[0, 2)$. So we have

$$\aleph = \overline{\overline{\mathcal{C}}} \le \overline{\overline{\mathcal{B} \cup \mathcal{C}}} \le \overline{\overline{[0, 2)}} = \aleph.$$

□

Remark. It is clear that

$$1 < 2 < \cdots < n < n+1 < \cdots < \aleph_0 < 2^{\aleph_0} < 2^{2^{\aleph_0}} < \cdots .$$

Therefore, **the set with maximal cardinality does not exist**.

Example. $\overline{\overline{R[a,b]}} > \overline{\overline{C[a,b]}} = \aleph$.

Proof. Let $\{r_1, \cdots, r_n, \cdots\} = [a,b] \cap \mathbb{Q}$, then there is a one-to-one mapping between $f \in C[a,b]$ to $\mathbb{R}^\infty$ by

$$f \mapsto \{f(r_1), f(r_2), \cdots, f(r_n), \cdots\} \in \mathbb{R}^\infty.$$

Therefore $\overline{\overline{C[a,b]}} \leq \overline{\overline{\mathbb{R}^\infty}} = \aleph$. On the other hand, any of constant functions belongs to $C[a,b]$. Notice that the set, K, of all constant functions has cardinality $\aleph$, so that $\mathbb{R}^\infty$ is equivalent to K, a subset of $C[a,b]$. This implies that $\overline{\overline{\mathbb{R}^\infty}} \leq \overline{\overline{C[a,b]}}$. Therefore, $\overline{\overline{C[a,b]}} = \aleph$ by Bernstein's theorem.

On the other hand, for every $A \in \mathcal{P}([0,1])$, define $\varphi(A) := \chi_A$. Then φ is an injective from $\mathcal{P}([0,1])$ to $R[a,b]$, so that

$$2^\aleph = \overline{\overline{\mathcal{P}([0,1])}} \leq \overline{\overline{R[a,b]}}.$$

Also, for every $f \in R[a,b]$, define

$$\phi(f) := \{(x, f(x)) : x \in [0,1]\} \subset [0,1] \times \mathbb{R}.$$

Then ϕ is an injective from $R[a,b]$ to $\mathcal{P}([0,1] \times \mathbb{R})$, so that

$$\overline{\overline{R[a,b]}} \leq \mathcal{P}([0,1] \times \mathbb{R}) = 2^\aleph,$$

where the equality follows from the fact that $\overline{\overline{[0,1] \times \mathbb{R}}} = \aleph$ (WHY). □

§3. Equivalence relations, orderings and axiom of choice

Equivalence relations. Let A be a nonempty set. A relation, $\sim$, on A is said to be an **equivalence relation** if for all $x, y, z \in A$,

- $x \sim x$ (reflexive)
- $x \sim y$ implies that $y \sim x$ (symmetric)
- $x \sim y$ and $y \sim z$ implies that $x \sim z$ (transitive)

Equivalence classes. Let A be a nonempty set and $\sim$ an equivalence relation on A. For each $x \in A$, define $E_x := \{y \in A : y \sim x\}$. And let $B := \{E_x : x \in A\}$. Each member of B is called an **equivalence class** of A under $\sim$.

Properties. Let A be a nonempty set and $\sim$ an equivalence relation on A. Then

- for each $x, y \in A$, either $E_x \cap E_y = \emptyset$ or $E_x = E_y$;
- $A = \bigcup_{x \in A} E_x$;
- $\sim$ partitions A into disjoint equivalence classes, that is, A is a disjoint union of the equivalence classes under $\sim$.

Ordering relations.

Definition. (**Partial ordering**) Let A be a set. A binary relation defined between certain pairs (x, y) of elements of A, expressed by $x \prec y$, is called a **partial ordering** on A if for all $x, y, z \in A$,

(i) (**reflexivity**) $x \prec x$,

(ii) (**antisymmetry**) $x \prec y$ and $y \prec x$ implies that $x = y$,

(iii) (**transitivity**) $x \prec y$ and $y \prec z$ implies that $x \prec z$.

In addition, a set endowed with a partial ordering is called a **partially ordered set**.

Definition. (**Total ordering**) A partial ordering $\prec$ is called a **total ordering** if additionally,

(iv) for every pair (x, y) in A, either $x \prec y$ or $y \prec x$.

A set endowed with a total ordering is called a **totally ordered set**.

Examples.

(1) $(\mathcal{P}(A), \subset)$ is a partially ordered set.
(2) $(\mathbb{R}, \leq)$ is a totally ordered set.
(3) Let G be a group. Let S be the set of subgroups with the relation that $H \prec H'$ if H is a subgroup of H'. Then $(S, \prec)$ is a partially ordered set. Given two subgroups, H, H' of G, we do not necessarily have $H \prec H'$ or $H' \prec H$.

Induced orderings. Let $(A, \prec)$ be a partially ordered set, and B a subset of A. We can define a partial ordering on B by defining that $x \prec y$ for $x, y \in B$ to hold if and only if $x \prec y$ in A. We shall say that it is the partial ordering on B induced by the ordering on A.

Upper bounds and maximal elements. Let $(A, \prec)$ be a partially ordered set, and B a subset of A. An **upper bound** of B (in A) is an element $x \in A$ such that $y \prec x$ for all $y \in B$. By a **maximal element** m of A one means an element of A such that if $x \in A$ and $m \prec x$, then $m = x$. In addition, we call $s \in A$ is a **minimal element** of A if $s \prec x$ for all comparable $x \in A$.

Zermelo's axiom of choice.

Axiom of choice. Suppose that $\mathcal{C}$ is a collection of nonempty sets. Then there exists a mapping $\varphi : \mathcal{C} \to \bigcup_{A\in\mathcal{C}} A$ such that $\varphi(A) \in A$ for each $A \in \mathcal{C}$.

Remarks. (1) An equivalent statement of the Zermelo's axiom of choice is: suppose that $\mathcal{C}$ is a collection of nonempty sets, then there exists a set B such that $B \subset \bigcup_{A\in\mathcal{C}} A$ and $A \cap B$ has only one element for each $A \in \mathcal{C}$.

(2) Roughly speaking, the axiom of choice asserts that given a collection of nonempty sets, it is possible to select an element from each set in the collection.

(3) Although most mathematicians use the axiom of choice without hesitation, some employ it only when they cannot obtain a proof without it and others consider it unacceptable. In real analysis and functional analysis, we will accept the axiom of choice and apply it freely.

Zorn's lemma. (**Principle of transfinite induction**) Let A be a nonempty partially ordered set with the property that every totally ordered subset of A has an upper bound in X. Then A contains at least one maximal element.

Well ordering principle. Every set X has at least one well-ordering.

Trichotomy of cardinality. Let A and B be two sets. Then exactly one of the following three statements holds:

$$\overline{\overline{A}} < \overline{\overline{B}}, \quad \overline{\overline{A}} = \overline{\overline{B}}, \quad \overline{\overline{A}} > \overline{\overline{B}}.$$

Continuum hypothesis. For every infinite cardinality m, there is no cardinality n such that $m < n < 2^m$.

Theorem. The following five results are equivalent.

(1) Axiom of choice.

(2) Zorn's lemma.

(3) Well ordering principle.

(4) Trichotomy of cardinality.

(5) Continuum hypothesis.

Hilbert's 23 problems - 1900.

1. **Continuum hypothesis** (G. Cantor, 1878):

• The contributions of K. Gödel (1940) and P. Cohen (1963) showed that the hypothesis can neither be disproved nor be proved using the axioms of Zermelo-Fraenkel set theory, the standard foundation of modern mathematics, provided ZF set theory is consistent. However, there is **no consensus** on whether this is a solution to the problem.

• Recent work of W. H. Woodin (2010,2011) has raised "hope" that there is an imminent solution.

• Is the Continuum Hypothesis a definite mathematical problem? – S. Feferman (2011)

2. The compatibility of the arithmetical axioms.

3. The equality of the volumes of two tetrahedra of equal bases and equal altitudes.

4. Problem of the straight line as the shortest distance between two points.

5. Lie's concept of a continuous group of transformations without the assumption of the differentiability of the functions defining the group.

6. Mathematical treatment of the axioms of physics.

7. Irrationality and transcendence of certain numbers.

8. Problems of prime numbers.

9. Proof of the most general law of reciprocity in any number field.

10. Determination of the solvability of a Diophantine equation.

11. Quadratic forms with any algebraic numerical coefficients.

12. Extension of Kroneckers theorem on abelian fields to any algebraic realm or rationality.

13. Impossibility of the solution of the general equation of the 7th degree by means of functions of only two arguments.

14. Proof of the finiteness of certain complete systems of functions.

15. Rigorous foundation of Schuberts enumerative calculus.

16. Problem of the topology of algebraic curves and surfaces.

17. Expression of definite forms by squares.

18. Building up of space from congruent polyhedra.

19. Are the solutions of regular problems in the calculus of variations always necessarily analytic?

20. The general problem of boundary values.

21. Proof of the existence of linear differential equations having a prescribed monodromic group.

22. Uniformization of analytic relations by means of automorphic functions.

23. Further development of the methods of the calculus of variations.

§5. Point sets on the line

Intervals.

- $(a, b) := \{x : a < x < b\}$, $-\infty \leq a < b \leq \infty$
- $[a, b) := \{x : a \leq x < b\}$, $-\infty < a < b \leq \infty$
- $(a, b] := \{x : a < x \leq b\}$, $-\infty \leq a < b < \infty$
- $[a, b] := \{x : a \leq x \leq b\}$, $-\infty < a \leq b < \infty$

In particular, $[a, a] = \{a\}$.

Bounded sets.

Definition. Let A be a nonempty subset of $\mathbb{R}$. A real number c is called an **upper bound** (or **lower bound**) of A if $x \leq c$ (or $c \leq x$) for all $x \in A$. If a subset of $\mathbb{R}$ has an upper bound (or lower bound), then we say that it is **bounded above** (or **bounded below**). A set is called **bounded** if it is bounded both above and below.

Definition. A real number u is called a **least upper bound** or **supremum** of A if it is an upper bound of A and smaller than or equal to any other upper bound of A. We write u by

$$\sup A, \quad \sup_{x \in A} x, \text{ or } \sup\{x : x \in A\}.$$

Definition. A real number l is called a **greatest lower bound** or **infimum** of A if it is a lower bound of A and greater than or equal to any other lower bound of A. We write l by

$$\inf A, \quad \inf_{x \in A} x, \text{ or } \inf\{x : x \in A\}.$$

Remark. We define $\sup A := +\infty$ if A is not bounded above, and define $\inf A := -\infty$ if A is not bounded below; In addition, (whenever needed) we may define

$$\sup \emptyset := -\infty, \quad \inf \emptyset := +\infty.$$

Open sets.

Definition. A subset $O \subset \mathbb{R}$ is said to be an **open set** if for each $x \in O$, there is an $r > 0$ such that $(x-r, x+r) \subset O$.

Remark. The open interval $(x-r, x+r)$, denoted by $N(x, r)$, is also called an r**-neighborhood of** x. Thus, A subset $O \subset \mathbb{R}$ is open if and only if for each point $x \in O$, there is an r_x-neighborhood $N(x, r_x) \subset O$.

Properties.

- $\emptyset$ and $\mathbb{R}$ are open sets.
- If A and B are open sets, then so is $A \cap B$. (**closed for finite intersection**)
- If $\{O_\alpha\}_{\alpha\in\Lambda}$ is a collection of open sets, then $\bigcup_{\alpha\in\Lambda} O_\alpha$ is open. (**closed for arbitrary union**)

Theorem. (**constructions of open sets**) Each open set O is countable union of disjoint open intervals. The representation is unique in the sense that if $\mathcal{C}$ and $\mathcal{D}$ are two pairwise disjoint collections of open intervals whose union is O, then $\mathcal{C} = \mathcal{D}$.

Proof. Let O be an open set. For $x \in O$ fixed, define

$$A_x := \{y : y < x \text{ and } (y, x) \subset O\},$$
$$B_x := \{z : z > x \text{ and } (x, z) \subset O\}.$$

The sets A_x and B_x are nonempty because O is open. Let $a_x := \inf A_x$ and $b_x := \sup B_x$. Then $a_x < x < b_x$ and $a_x, b_x \notin O$.

Set $I_x := (a_x, b_x)$ and note that $x \in I_x$. We claim that $I_x \subset O$. Indeed, let $u \in I_x$; then $a_x < u < b_x$. Thus, we can choose $y \in A_x$ and $z \in B_x$ such that $y < u < z$. If $u \leq x$, then $u \in (y, x] \subset O$ and, if $u > x$, then $u \in (x, z) \subset O$. Hence, $I_x \subset O$, so that $\bigcup_{x\in O} I_x \subset O$. On the other hand, as $x \in I_x$, $\bigcup_{x\in O} I_x \supset O$. Thus, by write $\mathcal{C} := \{I_x : x \in O\}$, we have $O = \bigcup_{A\in\mathcal{C}} A$.

Next we prove that $\mathcal{C}$ is countable. Indeed, it is easy to show that either $I_x \cap I_y = \emptyset$ or $I_x = I_y$. Let $A \in \mathcal{C}$. Because A is an open interval, we can, by the density of the rational numbers, select a rational number $r_A \in A$. Then define $\varphi : \mathcal{C} \to Q$ by $\varphi(A) = r_A$. Note that φ is one-to-one because $\mathcal{C}$ is pairwise disjoint. Thus, $\mathcal{C}$ is equivalent to a subset of $\mathbb{Q}$ and, consequently, is countable.

Finally, we prove the uniqueness of the representation. Let $\mathcal{D}$ be a pairwise disjoint collection of open intervals whose union is O. For each interval $(a, b) \in \mathcal{D}$, we claim that $a, b \notin O$. Indeed, suppose to the contrary that $a \in O$. Then there is another open interval $(c, d) \in \mathcal{D}$ such that $a \in (c, d)$, so that $(a, b) \cap (c, d) \neq \emptyset$. But this is impossible because $\mathcal{D}$ is a pairwise disjoint collection of open intervals. Thus, $a \notin O$ and, similarly, $b \notin O$. Therefore, for each $x \in O$, there is a unique open interval $I'_x \in \mathcal{D}$ such that $x \in I'_x$. Further on, we can show that $I'_x = I_x$. This proves that $\mathcal{C} = \mathcal{D}$. □

Limit points.

Definition. Let $A \subset \mathbb{R}$. A real number x_0 is called a **limit point** of A if

$$(N(x_0, \epsilon) \setminus \{x_0\}) \cap A \neq \emptyset$$

for each $\epsilon > 0$. The set of all limit points of A is denoted by A'.

Proposition. Let $A \subset \mathbb{R}$ and $x_0 \in \mathbb{R}$. The following statements are equivalent.

(1) x_0 is a limit point of A.

(2) There exists a sequence $\{x_n\} \subset A$, with $x_n \neq x_0$ for $n = 1, 2, \cdots$, such that $x_n \to x_0$ as $n \to \infty$.

(3) For each $\epsilon > 0$, the set $A \cap N(x_0, \epsilon)$ is infinite set.

Isolated points.

Definition. Let $A \subset \mathbb{R}$. A point $x_0 \in A$ is called an **isolated point** if there is an r-neighborhood $N(x_0, r)$ such that

$$(N(x_0, r) \setminus \{x_0\}) \cap A = \emptyset.$$

In addition, a set $A \subset \mathbb{R}$ is called an **isolated set** if each point of A is isolated.

Proposition. (i) A set A is isolated set if and only if $A \cap A' = \emptyset$.

(ii) A point x is not isolated if and only if x is a limit point.

Examples. (1) A set with no limit point: $\{1, 2, 3, \cdots, n\}$.

(2) A set not containing its limit points:

$$E = \left\{1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{n}, \cdots\right\}.$$

Clearly, $E' = \{0\}$.

(3) A set containing some of its limit points:

$$E = \mathbb{Q} \cap [0, 1].$$

Clearly, $E' = [0, 1]$.

(4) A set containing all its limit points: $E = [0, 1] = E'$.

Closed sets.

Definition. A subset $F \subset \mathbb{R}$ is said to be a **closed set** if $F' \subset F$, that is, if F contains all its limit points.

Proposition. Let $F \subset \mathbb{R}$. The following statements are equivalent.

(1) F is closed.

(2) $F^c = \mathbb{R} \setminus F$ is open.

(3) For each sequence $\{x_n\} \subset F$ with $\lim_{n\to\infty} x_n = x \in \mathbb{R}$, then $x \in F$.

Properties.

- $\emptyset$ and $\mathbb{R}$ are closed sets.
- If A and B are closed sets, then so is $A \cup B$. (**closed for finite union**)
- If $\{F_\alpha\}_{\alpha\in\Lambda}$ is a collection of closed sets, then $\bigcap_{\alpha\in\Lambda} F_\alpha$ is closed. (**closed for arbitrary intersection**)

Remark. If a subset of $\mathbb{R}$ is both open and closed, then it is either $\emptyset$ or $\mathbb{R}$.

Closures of sets.

Definition. The **closure of a subset** $F \subset \mathbb{R}$, denoted by $\overline{F}$, is defined by $\overline{F} := F \cup F'$.

Properties.

- The closure of a set $A \subset \mathbb{R}$ is closed.
- Let $A \subset \mathbb{R}$. A point $x \in \overline{A}$ if and only if $A \cap N(x, \epsilon) \neq \emptyset$ for each $\epsilon > 0$.
- A subset $A \subset \mathbb{R}$ is closed if and only if $A = \overline{A}$.

Dense sets.

Definition. Let $X \subset \mathbb{R}$. A subset $A \subset X$ is said to be **dense in** X if $\overline{A} = X$.

Proposition. Let $A \subset X$. The following statements are equivalent.

(i) A is dense in X.

(ii) Every point of X is a limit point of A.

(iii) $N(x, r) \cap A \neq \emptyset$ for all $x \in X$ and $r > 0$.

Perfect sets.

Definition. A subset $A \subset \mathbb{R}$ is said to be **perfect set** if $A = A'$, that is, each point of A is the limit point.

Remark. Clearly, a subset $A \subset \mathbb{R}$ is a perfect set if and only if it is a closed set with no isolated point.

Examples. $\emptyset, \mathbb{R}, [a, b](a < b)$.

Nowhere dense sets.

Definition. A subset $A \subset \mathbb{R}$ is called **nowhere dense** if $\overline{A}$ does not contain any open interval.

The Cantor set.

Base-p expansions. Let p be an integer greater than 1. Then for each $x \in [0, 1]$, there is a sequence $\{a_n\}$ of integers such that $0 \leq a_n \leq p - 1$ for all n and

$$x = \sum_{n=1}^{\infty} \frac{a_n}{p^n} = \frac{a_1}{p} + \frac{a_2}{p^2} + \frac{a_3}{p^3} + \cdots . \tag{0.2}$$

The sequence $\{a_n\}$ is unique unless $x \neq 1$ and is of the form $\frac{q}{p^m}$ for some $q, m \in \mathbb{N}$, in which case there are exactly two such sequences, one having only finitely many nonzero terms and the other having only finitely many terms different from $p - 1$.

Also, we use the notation

$$x = 0.a_1a_2a_3 \cdots \quad (p)$$

as a shorthand for the expansion (0.2).

Examples.

- For each integer $p \geq 2$, we have

$$0 = 0.000 \cdots \quad (p)$$

and

$$1 = 0.(p - 1)(p - 1)(p - 1) \cdots \quad (p)$$

• The number 1/2 has, respectively, the binary (p=2), ternary (p=3) and decimal (p=10) expansions given by

$$\begin{aligned} 1/2 =& 0.1000\cdots \quad (2) \\ =& 0.1111\cdots \quad (3) \\ =& 0.5000\cdots \quad (10). \end{aligned}$$

Notice that $1/2 = \frac{1}{2^1}$ with $m = 1$ and $q = 1$, and $1/2 = \frac{5}{10^1}$ with $m = 1$ and $q = 5$. 1/2 also has a second binary expansion and decimal expansion:

$$\begin{aligned} 1/2 =& 0.0111\cdots \quad (2) \\ =& 0.4999\cdots \quad (10). \end{aligned}$$

However, the ternary expansion of 1/2 is unique.

The Cantor set.

In mathematics, the Cantor set, introduced by German mathematician Georg Cantor in 1883 (but discovered in 1875 by Henry John Stephen Smith), is a set of points lying on a single line segment that has a number of remarkable and deep properties. Through consideration of it, Cantor and others helped lay the foundations of modern general topology. Although Cantor himself defined the set in a general, abstract way, the most common modern construction is the Cantor ternary set, built by removing the middle thirds of a line segment. Cantor himself only mentioned the ternary construction in passing, as an example of a more general idea, that of a perfect set that is nowhere dense.

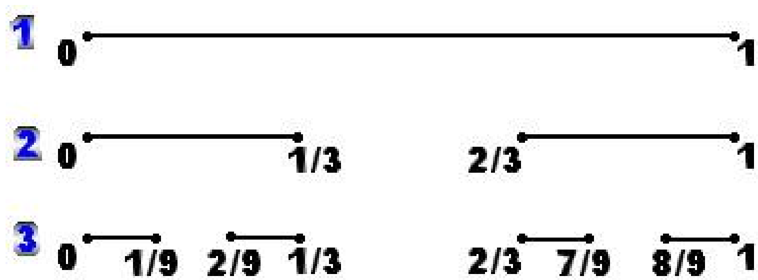
==Construction of the ternary set==

The Cantor ternary set is created by repeatedly deleting the open middle thirds of a set of line segments. One starts by deleting the open middle third $(\frac{1}{3}, \frac{2}{3})$ from the interval $[0, 1]$, leaving two line segments: $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. Next, the open middle third of each of these remaining segments is deleted, leaving four line segments: $[0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{7}{9}, 1]$. This process is continued ad infinitum, where the "n"th set is

$$\frac{C_{n-1}}{3} \cup \left(\frac{2}{3} + \frac{C_{n-1}}{3}\right).$$

The Cantor ternary set contains all points in the interval $[0, 1]$ that are not deleted at any step in this infinite process.

The first three steps of this process are illustrated below.



== Composition ==

Since the Cantor set is defined as the set of points not excluded, the proportion (i.e., Lebesgue measure) of the unit interval remaining can be found by total length removed. This total is the geometric progression

$$\sum_{n=0}^{\infty} \frac{2^n}{3^{n+1}} = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \frac{8}{81} + \cdots = \frac{1}{3}\left(\frac{1}{1-\frac{2}{3}}\right) = 1.$$

So that the proportion left is $1 - 1 = 0$.

This calculation shows that the Cantor set cannot contain any interval of non-zero length. In fact, it may seem surprising that there should be anything left – after all, the sum of the lengths of the removed intervals is equal to the length of the original interval. However, a closer look at the process reveals that there must be something left, since removing the "middle third" of each interval involved removing open sets (sets that do not include their endpoints). So removing the line segment $(\frac{1}{3}, \frac{2}{3})$ from the original interval $[0, 1]$ leaves behind the points $\frac{1}{3}$ and $\frac{2}{3}$. Subsequent steps do not remove these (or other) endpoints, since the intervals removed are always internal to the intervals remaining. So the Cantor set is not empty, and in fact contains an infinite number of points.

It may appear that "only" the endpoints are left, but that is not the case either. The number $1/4$, for example is in the bottom third, so it is not removed at the first step, and is in the top third of the bottom third, and is in the bottom third of "that", and in the "top" third of "that", and so on ad infinitum; alternating between top and bottom thirds. Since it is never in one of the middle thirds, it is never removed, and yet it is also not one of the endpoints of any middle third. The number $3/10$ is also in the Cantor set and is not an endpoint.

In the sense of cardinality, "most" members of the Cantor set are not endpoints of deleted intervals.

=== Cardinality ===

It can be shown that there are as many points left behind in this process as there were that were removed, and that therefore, the Cantor set is uncountable. To see this, we show that there is a function f from the Cantor set G to the closed interval $[0, 1]$ that is surjective (i.e. f maps from G onto $[0, 1]$) so that the cardinality of G is no less than that of $[0, 1]$. Since G is a subset of $[0, 1]$, its cardinality is also no greater, so the two cardinalities must in fact be equal.

To construct this function, consider the points in the $[0, 1]$ interval in terms of base 3 (or ternary numeral system) notation. As showed above,

$$\begin{aligned} 1/3 =& 0.1000\cdots \quad (3), \\ 2/3 =& 0.2000\cdots \quad (3) \\ =& 0.1222\cdots \quad (3). \end{aligned}$$

Thus, the middle third (to be removed) contains the numbers with ternary numerals of the form

$$0.1a_2a_3a_4\cdots \quad (3),$$

where $a_k \in \{0, 1, 2\}$ and not all the a_ks are 0 and not all are 2. So the numbers remaining after the first step consists of

- Numbers of the form $0.0a_2a_3a_4\cdots \quad (3), \quad$ (points in the interval $[0, 1/3)$)
- $1/3 = 0.1000\cdots \quad (3) = 0.0222\cdots \quad (3)$,
- $2/3 = 0.1222\cdots \quad (3) = 0.2000\cdots \quad (3)$,
- Numbers of the form $0.2a_2a_3a_4\cdots \quad (3), \quad$ (points in the interval $(2/3, 1]$)

where $a_k \in \{0, 1, 2\}$ for $k = 2, 3, \cdots$. All the points remained can be restated as those numbers with a ternary numeral

$$0.0a_2a_3a_4\cdots \quad (3)$$

or

$$0.2a_2a_3a_4\cdots \quad (3)$$

with $a_k \in \{0, 1, 2\}$ for $k = 2, 3, \cdots$.

The second step removes numbers of the form

$$0.01a_3a_4\cdots \quad (3)$$

and

$$0.21a_3a_4\cdots \quad (3),$$

and (with appropriate care for the endpoints) it can be concluded that the remaining numbers are those with a ternary numeral whose first "two" digits are not 1, that is, the numbers of the forms

$$0.00a_3a_4a_5\cdots \quad (3), \quad 0.02a_3a_4a_5\cdots \quad (3)$$

or

$$0.20a_3a_4a_5\cdots \quad (3), \quad 0.22a_3a_4a_5\cdots \quad (3),$$

where $a_k \in \{0, 1, 2\}$ for $k = 3, 4, \cdots$.

Continuing in this way, for a number not to be excluded at step "n", it must have a ternary representation whose "n"th digit is not 1. For a number to be in the Cantor set, it must not be excluded at any step, it must have a numeral consisting entirely of 0s and 2s. Thus, we obtain

Property. The Cantor set G consists of all numbers in $[0, 1]$ that have a ternary expansion without the digit 1.

The Cantor function Let $x \in G$. By the property above, x has a (unique) ternary expansion without the digit 1, say,

$$x = 0.(2c_1)(2c_2)(2c_3)\cdots \quad (3),$$

where $c_k \in \{0, 1\}$ for each $k \in \mathbb{N}$. We define **the Cantor function** $f : G \to [0, 1]$ by

$$f(x) := 0.c_1c_2c_3\cdots \quad (3).$$

Property. The Cantor function is a surjective from G to $[0, 1]$, that is, the range of f is $[0, 1]$. Hence, $\overline{\overline{G}} = \aleph$.

Proof. Let $y \in [0, 1]$. Rewrite it by the binary notation

$$y = 0.d_1d_2d_3\cdots \quad (2),$$

where $d_k \in \{0, 1\}$ for each $k \in \mathbb{N}$. Let

$$x = 0.(2d_1)(2d_2)(2d_3)\cdots \quad (3).$$

Then $x \in G$ and $f(x) = y$. That is, each point in $[0, 1]$ has an inverse image under f. This proves $R(f) = [0, 1]$.

Thus, $\overline{\overline{G}} \geq \overline{\overline{[0,1]}}$. Clearly, $\overline{\overline{G}} \leq \overline{\overline{[0,1]}}$ because $G \subset [0,1]$. Therefore, $\overline{\overline{G}} = \overline{\overline{[0,1]}} = \aleph$. $\square$

Remark. The set of endpoints of the removed intervals is countable, so there must be uncountably many numbers in the Cantor set which are not interval endpoints. As noted above, one example of such a number is $1/4$, which can be written as

$$0.02020202020\cdots \quad (3)$$

in ternary notation.

The Cantor set contains as many points as the interval from which it is taken, yet itself contains no interval. (Actually, the irrational numbers have the same property, but the Cantor set has the additional property of being closed, so it is not even dense in any interval, unlike the irrational numbers, which are dense everywhere.)

Properties.

- G is nowhere dense.
- G is perfect.
- $\overline{\overline{G}} = \aleph$.

Remark. As we have just seen, the complement in $[0,1]$ of the Cantor set, is disjoint union of open intervals, the sum of whose lengths is 1. But the length of $[0,1]$ is also 1. Thus, from the point of view of length, the Cantor set appears to be "**small**". On the other hand, G is uncountable, so that from a cardinality point of view, the Cantor set is "**large**". These, among other properties of the Cantor set, make it useful for illustrating many subtle concepts (for example, the "measure" so called).

Borel sets.

• G_δ **set**: the intersection of countable open subsets of $\mathbb{R}$.

• F_σ **set**: the union of countable closed subsets of $\mathbb{R}$.

• **Borel σ-algebra** on $\mathbb{R}$, denoted by $\mathscr{B}(\mathbb{R})$: the σ-algebra generated by the collection of all open sets on $\mathbb{R}$. Each element of $\mathscr{B}(\mathbb{R})$ is called a **Borel set**.

Finally, we point out that the complement of a G_δ set is an F_σ set, and the complement of an F_σ set is a G_δ set. In addition, G_δ sets and F_σ sets are necessarily Borel sets.

Examples of continuity.

(i) A function defined on $[0,1]$ which is continuous nowhere:

$$D(x) = \begin{cases} 1, & x \in [0,1] \cap \mathbb{Q}; \\ 0, & x \in [0,1] \setminus \mathbb{Q}. \end{cases}$$

(ii) A function defined on $[0,1]$ which is continuous only on one point:

$$f(x) = \begin{cases} x, & x \in [0,1] \cap \mathbb{Q}; \\ -x, & x \in [0,1] \setminus \mathbb{Q}. \end{cases}$$

(iii) A function defined on $\mathbb{R}$ which is continuous only on finite points $x_1, x_2, \cdots, x_n \in \mathbb{R}$:

$$f(x) = \begin{cases} (x-x_1)(x-x_2)\cdots(x-x_n), & x \in \mathbb{Q}; \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

(iv) A function defined on $\mathbb{R}$ which is continuous only on $\mathbb{Z}$ (countable set):

$$f(x) = \begin{cases} \sin \pi x, & x \in \mathbb{Q}; \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

(v) A function defined on $[0,1]$ which is continuous only on an uncountable set (with Lebesgue measure > 0):

$$R(x) = \begin{cases} \frac{1}{q}, & x = \frac{p}{q} \quad (p, q \in \mathbb{Z}, (p,q) = 1); \\ 0, & x \in \{0,1\} \cup ([0,1] \setminus \mathbb{Q}). \end{cases}$$

(vi) A function defined on $[0,1]$ which is continuous only on $[0,1] \setminus G$:

$$f(x) = \begin{cases} 1, & x \in G; \\ 0, & x \in [0,1] \setminus G. \end{cases}$$

End of Chapter I

• Could you give two concrete sequences of sets which are not convergent?

• Could you give at least three distinct proofs for the statement

$$\aleph_0 < \aleph?$$

• Let $a \in \mathbb{R}$ and let $\mathcal{A} := \{A \subset \mathbb{R} : a \in A\}$. $\overline{\overline{\mathcal{A}}} = ?$

• Let $A^n := A \times A \times \cdots \times A$ and $A^\infty := A \times A \times \cdots$. Recall that

$$\overline{\overline{A}} = \overline{\overline{A^n}} = \overline{\overline{A^\infty}}$$

whenever $\overline{\overline{A}} = \aleph$. What about the conclusion if $\overline{\overline{A}} = \aleph_0$?

• In the classical formula

$$\lim_{n\to\infty} 2^n = \infty,$$

it is clear that the first infinity is exactly $\aleph_0$. What about the last infinity?

• Could you give two concrete relations in mathematics which do not satisfy the transitivity?

• Recall that there is a function $f \in R[0,1]$ such that f is continuous on $[0,1] \setminus \mathbb{Q}$ while discontinuous on $[0,1] \cap \mathbb{Q}$, for example, the **Riemann function**. Could you construct another concrete function sharing this interesting property.

• Could you construct two concrete functions $f, g \in R[0,1]$ such that the discontinuities of f are just the continuities of g?

• Could you construct two discontinuous functions $f, g \in R[0,1]$ such that the discontinuities of f are just the continuities of g?

• Could you construct a concrete function $g \in R[0,1]$ such that g is discontinuous on $[0,1] \setminus \mathbb{Q}$ while continuous on $[0,1] \cap \mathbb{Q}$?

Chapter 2: Theory of measures

§1. Lebesgue measure

Lengths of intervals: $l(I)$

Definition. (**Lengths of bounded intervals**) Let $a, b \in \mathbb{R}$ with $a < b$, and let I be a bounded interval, i.e., $I = (a, b)$, or $[a, b)$, or $(a, b]$, or $[a, b]$. The **length** $l(I)$ of I is defined by

$$l(I) := b - a.$$

In particular, we define $l(\{a\}) := 0$ as well as $l(\emptyset) := 0$.

Definition. (**Lengths of unbounded intervals**) Let $a, b \in \mathbb{R}$, and let I be an unbounded interval, i.e., $I = (a, +\infty)$, or $[a, +\infty)$, or $(-\infty, b]$, or $(-\infty, b)$, or $(-\infty, +\infty)$. The **length** $l(I)$ of I is defined by

$$l(I) := +\infty.$$

Definition. (**Lengths of finitely many unions of intervals**) Let $I_1, I_2, \cdots, I_n$ be disjoint intervals. The **length** of $I = \bigcup_{k=1}^n I_k$ is defined by

$$l(I) := \sum_{k=1}^{n} l(I_k).$$

Properties of lengths.

(i) (**Nonnegativity**) $l(I) \geq 0$ for each interval I.

(ii) (**Monotonicity**) $l(I_1) \leq l(I_2)$ for intervals $I_1 \subset I_2$.

(iii) (**Translation invariance**) $l(I) = l(x + I)$ for every real number x and interval I, where $x + I := \{x + y : y \in I\}$.

(iv) (**Finite subadditivity**) Let $I_1, I_2, \cdots, I_n$ be intervals. Then

$$l\Big(\bigcup_{k=1}^{n} I_k\Big) \le \sum_{k=1}^{n} l(I_k).$$

(v) (**Continuity**) Let $\{I_n\}$ be a sequence of intervals.

(a) If $I_1 \subset I_2 \subset \cdots$, then

$$l\big(\lim_{n\to\infty} I_n\big) = \lim_{n\to\infty} l(I_n).$$

(b) If $I_1 \supset I_2 \supset \cdots$ and $l(I_1) < +\infty$, then

$$l\big(\lim_{n\to\infty} I_n\big) = \lim_{n\to\infty} l(I_n).$$

Proof. The statements (i), (ii) and (iii) are direct consequences of the definition of length.

(iv) It suffices to prove the conclusion for that $I_1, \cdots, I_n$ are all bounded **open** intervals. Clearly, the case $n = 1$ is trivial. Suppose inductively that the conclusion holds for $n = k$. Then, for $n = k + 1$,

(a) if $I_1, \cdots, I_k, I_{k+1}$ are pairwise disjoint, then

$$l\Big(\bigcup_{i=1}^{k+1} I_i\Big) = \sum_{i=1}^{k+1} l(I_i)$$

by the definition of length of finitely many unions of intervals;

(b) otherwise, there are two intervals, I_{n_1} and I_{n_2} such that $I_{n_1} \cap I_{n_2} \neq \emptyset$. By writing I as the construction interval of $I_{n_1} \cup I_{n_2}$, we have $l(I) \le l(I_{n_1}) + l(I_{n_2})$, so that

$$\begin{aligned} l\Big(\bigcup_{i=1}^{k+1} I_i\Big) &= I \cup \Big(\bigcup_{i\neq n_1, n_2} I_i\Big) \\ &\le l(I) + \sum_{i\neq n_1, n_2} l(I_i) \le \sum_{i=1}^{k+1} l(I_i). \end{aligned}$$

Thus, we obtain the desired conclusion by the principle of induction.

(v) It sufficient to prove the conclusion for closed intervals. Let $I_n = [a_n, b_n]$ for $n \in \mathbb{N}$.

(a) Suppose that

$$\cdots \leq a_2 \leq a_1 < b_1 \leq b_2 \leq \cdots .$$

Clearly, $\lim_{n\to\infty} I_n = (a, b)$ with

$$-\infty \leq a := \lim_{n\to\infty} a_n, \quad b := \lim_{n\to\infty} b_n \leq +\infty,$$

and

$$\lim_{n\to\infty} l(I_n) = \lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = l\big(\lim_{n\to\infty} I_n\big).$$

(b) Suppose that

$$a_1 \leq \cdots \leq a_n \leq a_{n+1} \leq \cdots \leq b_{n+1} \leq b_n \leq \cdots \leq b_1.$$

Notice that I_1 is bounded, so that $\{a_n\}$ and $\{b_n\}$ are convergent by the monotone convergence criteria. Denote $a := \lim_{n\to\infty} a_n$ and $b := \lim_{n\to\infty} b_n$. Then

$$\lim_{n\to\infty} l(I_n) = \lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = l\big(\lim_{n\to\infty} I_n\big).$$

$\square$

Finally, we give a general monotonicity of the length of intervals.

Theorem. (**General monotonicity**) Let $I_1, \cdots, I_n$ be pairwise disjoint intervals, and also, let $J_1, \cdots, J_m$ be pairwise disjoint intervals. Then

$$\Big(\bigcup_{k=1}^{n} I_k\Big) \subset \bigcup_{k=1}^{m} J_k \quad \Rightarrow \quad l\Big(\bigcup_{k=1}^{n} I_k\Big) \leq l\Big(\bigcup_{k=1}^{m} J_k\Big).$$

Proof. It suffices to consider that all intervals above are open intervals. Further on, suppose that $\left(\bigcup_{k=1}^n I_k\right) \subset \bigcup_{k=1}^m J_k$. Clearly, for each $I_s \in \{I_1, \cdots, I_n\}$, there is a $J_k \in \{J_1, \cdots, J_m\}$ such that $I_s \subset J_k$. Fix $k \in \{1, \cdots m\}$ and write

$$E_k := \{I \in \{I_1, \cdots, I_n\} : I \subset J_k\}.$$

Then $l(\cup_{I \in E_k} I) \le l(J_k)$ (WHY) and

$$\bigcup_{k=1}^n I_k = \bigcup_{k=1}^m \Big(\bigcup_{I \in E_k} I\Big).$$

Therefore, from the finite additivity it follows that

$$\begin{aligned} l\Big(\bigcup_{k=1}^n I_k\Big) &= \sum_{k=1}^n l(I_k) = \sum_{k=1}^m \Big(\sum_{I \in E_k} l(I)\Big) \\ &= \sum_{k=1}^m l\Big(\bigcup_{I \in E_k} I\Big) \le \sum_{k=1}^m l(J_k) = l\Big(\bigcup_{k=1}^m J_k\Big). \end{aligned}$$

□

By the decomposition of open sets, we have the following

Corollary. Let $I_1, \cdots, I_n$ and $J_1, \cdots, J_m$ be general intervals (not necessary to be pairwise disjoint). Then

$$\Big(\bigcup_{k=1}^n I_k\Big) \subset \bigcup_{k=1}^m J_k \quad \Rightarrow \quad l\Big(\bigcup_{k=1}^n I_k\Big) \le l\Big(\bigcup_{k=1}^m J_k\Big).$$

Problem. Under the same condition of above corollary, can we have the inequality $\sum_{k=1}^n l(I_k) \le \sum_{k=1}^m l(J_k)$?

Corollary. Let $I_1, \cdots, I_n$ be pairwise disjoint intervals, and let $J_1, \cdots, J_m$ be general intervals (not necessary to be pairwise disjoint). Then

$$\left(\bigcup_{k=1}^{n} I_k\right) \subset \bigcup_{k=1}^{m} J_k \quad \Rightarrow \quad \sum_{k=1}^{n} l(I_k) \leq \sum_{k=1}^{m} l(J_k).$$

Measures of open sets: $m(O)$

By the decomposition of open subsets of $\mathbb{R}$, an open subset $O \subset \mathbb{R}$ is the countable union of pairwise disjoint open intervals $I_i = (a_i, b_i)$ (**construction intervals** so called), i.e., $O = \bigcup_{i=1}^{\infty}(a_i, b_i)$, so we can define a "length" of O via the length of open interval.

Definition. Let $O \subset \mathbb{R}$ be open. We define its **measure** $m(O)$ by

$$m(O) := \sum_{i=1}^{\infty} l(I_i) = \sum_{i=1}^{\infty}(b_i - a_i),$$

with (a_i, b_i), $i = 1, 2, \cdots$, the construction intervals of O.

Remark. Since every open interval I is an open set, by the definition of measure of open sets, we have

$$\begin{aligned} m(I) =& m(I \cup \emptyset \cup \emptyset \cup \cdots) \\ =& l(I) + 0 + 0 + \cdots = l(I). \end{aligned}$$

Thus, m (as a function with domain $\{O \subset \mathbb{R} : O \text{ is open}\}$) is an extension of the length function l (with domain $\{I \subset \mathbb{R} : I \text{ is open interval}\}$).

Properties.

(i) (**Nonnegativity**) $m(O) \geq 0$ for each open set $O \subset \mathbb{R}$.

(ii) (**Monotonicity**) Let O_1 and O_2 be two open subsets of $\mathbb{R}$. If $O_1 \subset O_2$, then $m(O_1) \leq m(O_2)$.

(iii) (**Translation invariance**) $m(O) = m(x + O)$ for every real number x and open set $O \subset \mathbb{R}$, where $x + O := \{x + y : y \in O\}$.

(iv) (**Countable additivity**) Let $\{O_n\}$ be a sequence of pairwise disjoint open subsets of $\mathbb{R}$. Then

$$m\Big(\bigcup_{n=1}^{\infty} O_n\Big) = \sum_{n=1}^{\infty} m(O_n).$$

(v) (**Countable subadditivity for open intervals**) For a sequence $\{I_n\}$ of open intervals,

$$m\Big(\bigcup_{n=1}^{\infty} I_n\Big) \leq \sum_{n=1}^{\infty} l(I_n).$$

(vi) (**Countable subadditivity for open subsets**) Let $\{O_n\}$ be a sequence of open subsets of $\mathbb{R}$. Then

$$m\Big(\bigcup_{n=1}^{\infty} O_n\Big) \leq \sum_{n=1}^{\infty} m(O_n).$$

In particular, we have the following **finite subadditivity for open subsets**:

$$m\Big(\bigcup_{k=1}^{n} O_n\Big) \leq \sum_{k=1}^{n} m(O_n).$$

Proof. The statements (i) and (iii) are direct consequences of the definition of measure of open sets.

(ii) Suppose that $O_1 \subset O_2$. Let $I_1, I_2, \cdots$ and $J_1, J_2, \cdots$ be the construction intervals of O_1 and O_2, respectively. Note that, for each I_n, there is a J_k such that $I_n \subset J_k$. If there is a J_n with $l(J_n) = +\infty$, then the conclusion is trivial. So we suppose further that all the construction intervals of O_2 are bounded.

Let $m \in \mathbb{N}^+$ and $\epsilon > 0$. Write $I_n^\epsilon := [a_n + \frac{\epsilon}{2m}, b_n - \frac{\epsilon}{2m}]$ for $I_n := (a_n, b_n)$. Since $\bigcup_{k=1}^{m} I_k^\epsilon$ is a closed set, by Heine-Borel theorem there are $J_{n_1}, J_{n_2}, \cdots, J_{n_p}$ such that

$$\bigcup_{k=1}^{m} I_k^\epsilon \subset \bigcup_{k=1}^{p} J_{n_k}.$$

Thus, from the general monotonicity of the length of intervals it follows that

$$\begin{aligned}-\epsilon + \sum_{k=1}^{m} l(I_k) =& \sum_{k=1}^{n} l(I_k^\epsilon) = l\left(\bigcup_{k=1}^{m} I_k^\epsilon\right) \\ \leq& l\left(\bigcup_{k=1}^{p} J_{n_k}\right) = \sum_{k=1}^{p} l(J_{n_k}) \leq m(O_2).\end{aligned}$$

By letting $\epsilon \to 0$ in both end sides of above formula, we obtain

$$\sum_{k=1}^{m} l(I_k) \leq m(O_2), \quad m = 1, 2, \cdots.$$

Again, by letting $m \to \infty$ in both sides, we have

$$m(O_1) = \sum_{k=1}^{\infty} l(I_k) \leq m(O_2).$$

(iv) By the decomposition of the open subset of $\mathbb{R}$, each O_n is the countable union of pairwise disjoint open intervals of $\mathbb{R}$. Suppose that

$$O_n = \bigcup_{m=1}^{\infty} I_{nm}, \quad n = 1, 2, \cdots.$$

Since the union of countable many sets is also countable, we can rewrite

$$\bigcup_{n=1}^{\infty} O_n = \bigcup_{k=1}^{\infty} I_k,$$

where $I_k \in \{I_{nm} : n, m = 1, 2, \cdots\}$. Clearly, $I_i \cap I_j = \emptyset$ for $i \neq j$. Thus, by the definition of measure of open set, we have

$$\begin{aligned} m\Big(\bigcup_{n=1}^{\infty} O_n\Big) &= m\Big(\bigcup_{k=1}^{\infty} I_k\Big) \\ &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} l(I_{nm}) = \sum_{n=1}^{\infty} m(O_n) \end{aligned}$$

The statement (v) can be proved by analogous argument given in the proof of (ii). Indeed, the case $\sum_{k=1}^{\infty} l(I_k) = +\infty$ is trivial. Let $\sum_{k=1}^{\infty} l(I_k) < +\infty$ and let $\{J_j\}$ be a sequence of construction intervals of the open set $\bigcup_{k=1}^{\infty} I_k$. For each $n \in N$, noticing $\bigcup_{j=1}^{n} J_j \subset \bigcup_{k=1}^{\infty} I_k$, we have

$$\sum_{j=1}^{n} l(J_j) \leq \sum_{k=1}^{\infty} l(I_k).$$

(Indeed, write $J_j^{\epsilon} = [a_j + \frac{\epsilon}{2n}, b_j - \frac{\epsilon}{2n}]$ for $J_j = (a_j, b_j)$ ($j = 1, 2, \cdots, n$). Then $\bigcup_{k=1}^{n} J_k^{\epsilon}$ is a closed set, so that by Heine-Borel theorem there are $I_{k_1}, I_{k_2}, \cdots, I_{k_m}$ such that

$$\bigcup_{k=1}^{n} J_k^{\epsilon} \subset \bigcup_{i=1}^{m} I_{k_i}$$

Thus,

$$-\epsilon + \sum_{k=1}^{n}(b_k - a_k) = \sum_{k=1}^{n} l(J_k^{\epsilon}) \leq \sum_{i=1}^{m} l(I_{k_i}) \leq \sum_{k=1}^{\infty} l(I_k),$$

and by letting $\epsilon \to 0$ we obtain the inequality desired). Therefore, by letting $n \to \infty$ in both sides we obtain

$$m\Big(\bigcup_{n=1}^{\infty} I_n\Big) = m\Big(\bigcup_{k=1}^{\infty} J_k\Big) = \sum_{j=1}^{\infty} l(J_j) \leq \sum_{k=1}^{\infty} l(I_k).$$

(vi) is a direct consequence of (v). Indeed, let $I_{n1}, I_{n2}, \cdots$ be the construction intervals for each $n \in \mathbb{N}$. Notice that the countable union of countable sets is also countable. It follows from (v) that

$$\begin{aligned} m\left(\bigcup_{n=1}^{\infty} O_n\right) =& m\left(\bigcup_{n,k=1}^{\infty} I_{nk}\right) \le \sum_{n,k=1}^{\infty} l(I_{nk}) \\ &= \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} l(I_{nk}) = \sum_{n=1}^{\infty} m(O_n). \end{aligned}$$

□

Remark. **Countable additivity** (associated with the limit operation) is essential to the construction of measures.

Lebesgue outer measures: $m^*(A)$

Definition. Let $A \subset \mathbb{R}$. The **Lebesgue outer measure**, $m^*(A)$, of A is defined by

$$m^*(A) := \inf\left\{ \sum_n l(I_n) : I_n \text{ open intervals}, \bigcup_n I_n \supset A \right\},$$

where $\{I_n\}$ is either a **finite** or an **infinite** sequence.

Remark. For convenience, we call $\bigcup_n I_n$ an $\mathcal{L}$-cover of A if $A \subset \bigcup_n I_n$ with I_n, open intervals. The set of all $\mathcal{L}$-covers of A is denoted by $\mathcal{L}_A$. Thus,

$$m^*(A) = \inf\left\{ \sum_n l(I_n) : \bigcup_n I_n \in \mathcal{L}_A \right\}.$$

Remark. The outer measure m^* is an extended real-valued function with domain $\mathcal{P}(\mathbb{R})$, i.e.,

$$m^* : \mathcal{P}(\mathbb{R}) \to [0, \infty].$$

Example. $m^*(\emptyset) = m^*(\{x\}) = 0$. Indeed, for every $\epsilon > 0$, $I_\epsilon := (-\epsilon/2, \epsilon/2) \in \mathcal{L}_\emptyset$. Thus,

$$m^*(\emptyset) \le l(I_\epsilon) = \epsilon.$$

Applying $\epsilon \to 0$ in both sides yields $m^*(\emptyset) = 0$. By analogous argument, we can show that

$$m^*(\{x_1, x_2, \cdots, x_n\}) = 0$$

for arbitrary finite set $\{x_1, x_2, \cdots, x_n\}$.

Example. Let $x_1, x_2, \cdots \in \mathbb{R}$. Then

$$m^*(\{x_1, x_2, x_3, \cdots\}) = 0.$$

Indeed, let $\epsilon > 0$ and consider

$$L_n := \left(x_n - \frac{\epsilon}{2^{n+1}}, x_n + \frac{\epsilon}{2^{n+1}} \right), \quad n \in \mathbb{N}^+.$$

Clearly, $L := \bigcup_{n\in\mathbb{N}^+} L_n$ is an $\mathcal{L}$-cover of

$$X := \{x_1, x_2, x_3, \cdots\}.$$

Thus, by the definition of m^*, we have

$$m^*(X) \le \sum_{n=1}^{\infty} l(L_n) = \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon.$$

Finally, by letting $\epsilon \to 0$ we obtain $m^*(X) = 0$.

Basic properties of m^*.

(i) (**Nonnegativity**) $m^*(A) \ge 0$ for all $A \subset \mathbb{R}$.

(ii) (**Monotonicity**) $A \subset B \Rightarrow m^*(A) \le m^*(B)$.

(iii) (**Translation invariance**) $m^*(x + A) = m^*(A)$ for $A \subset \mathbb{R}$ and $x \in \mathbb{R}$, where $x + A := \{x + y : y \in A\}$.

(iv) (**Countable subadditivity**) For a sequence $\{A_n\}$ of subsets of $\mathbb{R}$,

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

In particular,

$$m^*(A \cup B) \le m^*(A) + m^*(B), \quad A, B \subset \mathbb{R}.$$

(v) $m^*(\overline{I}) = m^*(I) = l(I)$ for every interval I.

(vi) $m^*(O) = m(O)$ for each open subset $O \subset \mathbb{R}$.

Remark. From (v) and (vi), we can observe that m^* (as a function with domain $\mathcal{P}(\mathbb{R})$) is an extension of m (as a function defined on $\mathcal{O}$, the class of open subsets of $\mathbb{R}$) as well as l (as a function defined on $\mathcal{I}$, the class of intervals of $\mathbb{R}$), that is to say,

$$m^*|_{\mathcal{O}} = m, \quad m^*|_{\mathcal{I}} = l.$$

Proof. (i) is trivial. In addition, since $m(L) = m(x + L)$, the statement (iii) is a direct consequence of the fact that $L \in \mathcal{L}_A$ if and only if $x + L \in \mathcal{L}_{x+A}$.

(ii) Let $A \subset B$. Then every $\mathcal{L}$-cover of B is also an $\mathcal{L}$-cover of A. Thus, by the monotonicity of inf, we have

$$m^*(A) = \inf_{L \in \mathcal{L}_A} m(L) \leq \inf_{L \in \mathcal{L}_B} m(L) = m^*(B).$$

(iv) Let $\epsilon > 0$. By the definition of m^* (as an infimum), for each A_n there is an $\mathcal{L}$-cover $\bigcup_{k=1}^{\infty} I_{nk}$ of A_n such that

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

Notice that $\bigcup_{n=1}^{\infty} \left(\bigcup_{k=1}^{\infty} I_{nk} \right)$ is also an $\mathcal{L}$-cover of $\bigcup_{n=1}^{\infty} A_n$. Then the definition of m^* implies that

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

By letting $\epsilon \to 0$ we obtain the result desired, immediately.

(v) It sufficient to prove

$$l(I) \geq m^*(\overline{I}) \geq m^*(I) \geq l(I)$$

for bounded open interval $I = (a, b)$. First, for $\epsilon > 0$ small enough, we have $[a, b] \subset (a - \epsilon, b + \epsilon)$. Then by the definition of m^*, we have

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

By letting $\epsilon \to 0$ we obtain

$$m^*([a, b]) \leq b - a = l(a, b).$$

The second inequality desired is a direct consequence of the monotonicity of m^*.

Finally, let $\{I_n\}$ be a sequence of open intervals and $\bigcup_{n=1}^{\infty} I_n \supset I$. By the monotonicity and countable subadditivity of the length of interval, we have

$$l(I) \leq l\Bigg(\bigcup_{n=1}^{\infty} I_n\Bigg) \leq \sum_{n=1}^{\infty} l(I_n).$$

Taking the infimum with respect to such sequences $\{I_n\}$ yields

$$\begin{aligned} l(I) &\leq \inf\left\{\sum_{n=1}^{\infty} l(I_n) : I_n \text{ open intervals }, I \subset \bigcup_{n=1}^{\infty} I_n\right\} \\ &= m^*(l). \end{aligned}$$

This completes the proof of (v).

(vi) Let $\{I_n\}$ be the construction intervals of O. Clearly, $\bigcup_n I_n$ is an $\mathcal{L}$-cover of O. Notice that $\{I_n\}$ is pairwise disjoint, we then have

$$m(O) = \sum_{n=1}^{\infty} l(I_n) \geq m^*(O).$$

Conversely, let $\bigcup_{k=1}^{\infty} J_k$ be an $\mathcal{L}$-cover of O. From the monotonicity and countable additivity of m (for open sets), it follows that

$$m(O) \leq m\Bigg(\bigcup_{k=1}^{\infty} J_k\Bigg) \leq \sum_{k=1}^{\infty} m(J_k) = \sum_{k=1}^{\infty} l(J_k).$$

By taking the infimum with respect to such $\mathcal{L}$-covers of O we obtain

$$m(O) \leq m^*(O).$$

$\square$

Remark. It is necessary to point out that $m^*|_{\mathcal{O}}$, i.e, m, has the countable additivity, while m^* itself (defined on $\mathcal{P}(\mathbb{R})$) does not satisfy the countable additivity, even if the finite additivity, see Theorem 3.10 given in "A Course of Real Analysis", J. N. McDonald and N. A. Weiss, Page 117.

Theorem. Let $A \subset \mathbb{R}$. Then

$$m^*(A) = \inf\{m(O) : O \text{ is open and } O \supset A\}.$$

Proof. Write

$$a := \inf\{m(O) : O \text{ is open and } O \supset A\}.$$

Notice that $m^*(O) = m(O)$ for each open set O, and hence, by the monotonicity of m^* we have $m^*(A) \leq m(O)$ for $A \subset O$. By taking infimum over such open sets O we obtain

$$m^*(A) \leq a.$$

Conversely, let $\bigcup_n J_n$ be an $\mathcal{L}$-cover of A. It is clear that $\bigcup_n J_n$ is an open set containing A, so that

$$\sum_n l(J_n) = \sum_n m(J_n) \geq m\left(\bigcup_n J_n\right) \geq a.$$

By taking infimum over all $\mathcal{L}$-covers of A, we obtain

$$m^*(A) \geq a.$$

$\square$

Lebesgue measure: $m(E)$

Example. Let I be an interval. Then

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

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

Proof. Let $F \subset \mathbb{R}$. It suffices to show

$$m^*(F) \geq m^*(F \cap I) + m^*(F \cap I^c). \quad \text{(WHY)}$$

To this end, we take an $\mathcal{L}$-cover $\bigcup_{n=1}^{\infty} I_n$ of F. Then $I_n \cap I$ and $I_n \cap I^c$ are intervals or union of intervals (not necessarily to be open) and $I_n = (I_n \cap I) \cup (I_n \cap I^c)$ for all $n \in \mathbb{N}^+$. By the finite additivity of the length of interval, we have

$$l(I_n) = l(I_n \cap I) + l(I_n \cap I^c), \quad n \in \mathbb{N}^+.$$

Notice that

$$F \cap I \subset \bigcup_{n=1}^{\infty} (I_n \cap I), \quad F \cap I^c \subset \bigcup_{n=1}^{\infty} (I_n \cap I^c).$$

By the monotonicity and countable subadditivity of outer measures, we have

$$\begin{aligned}
&m^*(F \cap I) + m^*(F \cap I^c) \\
\leq &m^*\left(\bigcup_{n=1}^{\infty} (I_n \cap I)\right) + m^*\left(\bigcup_{n=1}^{\infty} (I_n \cap I^c)\right) \\
\leq &\sum_{n=1}^{\infty} m^*(I_n \cap I) + \sum_{n=1}^{\infty} m^*(I_n \cap I^c) \\
= &\sum_{n=1}^{\infty} l(I_n \cap I) + \sum_{n=1}^{\infty} l(I_n \cap I^c) = \sum_{n=1}^{\infty} l(I_n).
\end{aligned}$$

By taking infimum over all $\mathcal{L}$-covers of F, we obtain the desired inequality, immediately. $\square$

This example motivates us to define the Lebesgue measurable sets as follows.

Definition. (**Lebesgue measures**) A subset $E \subset \mathbb{R}$ is said to be a **Lebesgue measurable set** if, for each subset $F \subset \mathbb{R}$,

$$m^*(E \cap F) + m^*(E^c \cap F) = m^*(F). \tag{0.3}$$

The outer measure of a Lebesgue measurable set E is called the **Lebesgue measure** of E, and we write it by $m(E)$.

Remark. The equality (0.3) is called the **Carathéodory criterion**. Thus, a Lebesgue measurable set is such a set that the Carathéodory criterion is satisfied.

Remark. Let $\mathcal{M}$ be the class of all Lebesgue measurable sets. It is clear that m is an extended real-valued function from $\mathcal{M}$ to $[0, +\infty]$, i.e.,

$$m : \mathcal{M} \to [0, +\infty].$$

In addition, $m = m^*|_{\mathcal{M}}$.

Proposition. Let $E \subset \mathbb{R}$. The following statements are equivalent.

(i) E is Lebesgue measurable.

(ii) For every subsets $F \subset \mathbb{R}$ with $m^*(F) < +\infty$, there holds

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

Examples. (i) $\emptyset, \mathbb{R} \in \mathcal{M}$.

(ii) If $E \in \mathcal{M}$ then $E^c \in \mathcal{M}$.

(iii) $I \in \mathcal{M}$ for every interval I.

Problem. Whether $O \in \mathcal{M}$ for general open subset $O \subset \mathbb{R}$?

Proposition. (Properties of Lebesgue measurable sets)

(i) $\mathcal{M}$ is closed under the operations of union and intersection.

(i') $\mathcal{M}$ is closed under the operations of finitely many unions and intersections.

(ii) $\mathcal{M}$ is closed under the operations of countably many unions and intersections.

(iii) $\mathcal{M}$ is closed under the operation of limit.

Remark. The statement (i) implies that $\mathcal{M}$ is an algebra over $\mathbb{R}$. Further on, the statement (ii) implies that $\mathcal{M}$ is indeed a so called "σ-algebra" over $\mathbb{R}$.

Proof. (i) Since $A_1 \cap A_2 = (A_1^c \cup A_2^c)^c$, it suffices to show that $\mathcal{M}$ is closed under the operation of union. To this end, let $A_1, A_2 \in \mathcal{M}$. Then for each $F \subset \mathbb{R}$, we have

$$
\begin{aligned}
&m^*(F \cap (A_1 \cup A_2)) + m^*(F \cap (A_1 \cup A_2)^c)\\
=&m^*\Bigg((F \cap A_1) \cup (\underline{F \cap A_2})\Bigg) + m^*(F \cap A_1^c \cap A_2^c)\\
=&m^*\Bigg((F \cap A_1) \cup \underline{((F \cap A_2) \cap A_1) \cup ((F \cap A_2) \cap A_1^c)}\Bigg)\\
&+ m^*(F \cap A_1^c \cap A_2^c)\\
=&m^*\Bigg((F \cap A_1) \cup ((F \cap A_2) \cap A_1^c)\Bigg) + m^*(F \cap A_1^c \cap A_2^c)\\
\le&m^*(F \cap A_1) + m^*((F \cap A_1^c) \cap A_2) + m^*((F \cap A_1^c) \cap A_2^c)\\
=&m^*(F \cap A_1) + m^*(F \cap A_1^c) = m^*(F).
\end{aligned}
$$

This implies that $A_1 \cup A_2 \in \mathcal{M}$.

(ii) It suffices to show that $\mathcal{M}$ is closed under the operation of countably many unions. Let $\{A_n\}_{n=1}^{\infty} \subset \mathcal{M}$ and write $A := \bigcup_{n=1}^{\infty} A_n$. In order to show $A \in \mathcal{M}$, we consider a disjoint decomposition

$$A = \bigcup_{n=1}^{\infty} B_n$$

with $B_1 := A_1, B_2 := A_2 \setminus A_1, B_3 := A_3 \setminus (A_1 \cup A_2), \cdots$.

Let $F \subset \mathbb{R}$. For every $m \in \mathbb{N}^+$, it follows from (i') that $\bigcup_{n=1}^{m} B_n \in \mathcal{M}$, and hence,

$$\begin{aligned} m^*(F) \geq & m^*\left(\left(\bigcup_{n=1}^{m} B_n\right) \cap F\right) + m^*(A^c \cap F) \\ = & m^*\left(\bigcup_{n=1}^{m} (B_n \cap F)\right) + m^*(A^c \cap F) \\ = & \sum_{n=1}^{m} m^*(B_n \cap F) + m^*(A^c \cap F). \end{aligned}$$

Applying $m \to \infty$ yields

$$\begin{aligned} m^*(F) \geq & \sum_{n=1}^{\infty} m^*(B_n \cap F) + m^*(A^c \cap F) \\ \geq & m^*\left(\bigcup_{n=1}^{\infty} (B_n \cap F)\right) + m^*(A^c \cap F) \\ = & m^*(A \cap F) + m^*(A^c \cap F). \end{aligned}$$

This implies that $A \in \mathcal{M}$.

The statement (iii) is a direct consequence of (ii). Indeed, let $\{A_n\} \subset \mathcal{M}$ and $\lim_{n\to\infty} A_n = A$. From (ii) it follows that $\bigcup_{n\geq k} A_n := B_k \in \mathcal{M}$ for every $k \in \mathbb{N}^+$. Thus, by (ii) again, we have

$$A = \overline{\lim_{n\to\infty}} A_n = \bigcap_{k\geq 1} B_k \in \mathcal{M}.$$

$\square$

Problem. What about the Lebesgue measurability of the union $A \cup B$ for $A, B \notin \mathcal{M}$?

Proposition. (Properties of Lebesgue measure)

All the properties of Lebesgue outer measure hold for Lebesgue measure. Furthermore, the following statements hold.

(i) (**Subtractivity**) Let $A_1, A_2 \in \mathcal{M}$ and $A_1 \supset A_2$ with $m(A_2) < \infty$. Then

$$m(A_1 \setminus A_2) = m(A_1) - m(A_2).$$

(ii) (**Finite additivity**) Let $A_1, \cdots, A_n \in \mathcal{M}$ be pairwise disjoint. Then

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

(iii) (**Countable additivity**) Let $A_1, A_2, \cdots \in \mathcal{M}$ be pairwise disjoint, then

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

(iv) (**Continuity**) Let $\{A_n\} \subset \mathcal{M}$. Then

$$m(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} m(A_n)$$

whenever

(a) $A_1 \subset A_2 \subset \cdots$, or

(b) $A_1 \supset A_2 \supset \cdots$ and $m(A_1) < \infty$.

Proof. (i) It suffices to show that

$$m(A \cup B) = m(A) + m(B) \tag{0.4}$$

for disjoint $A, B \in \mathcal{M}$ (WHY). Indeed, let $A, B \in \mathcal{M}$ be disjoint. Since $A \in \mathcal{M}$ and $A^c \cap B = B$, we have

$$\begin{aligned} m(A \cup B) =& m(A \cap (A \cup B)) + m(A^c \cap (A \cup B)) \\ =& m(A) + m(A^c \cap B) = m(A) + m(B). \end{aligned}$$

The statement (ii) is also a direct consequence of (0.4) (WHY).

(iii) It suffices to show that

$$m\Big(\bigcup_{n=1}^{\infty} A_n\Big) \geq \sum_{n=1}^{\infty} m(A_n). \quad \text{(WHY)}$$

Clearly, from the finite additivity and monotonicity of m, it follows that

$$\sum_{k=1}^{n} m(A_k) = m\Big(\bigcup_{k=1}^{n} A_k\Big) \leq m\Big(\bigcup_{k=1}^{\infty} A_k\Big), \quad n \in \mathbb{N}^+.$$

By letting $n \to \infty$ in both sides, we obtain the desired inequality, immediately.

(iv) (a) The conclusion is trivial if there is an $n \in \mathbb{N}^+$ such that $m(A_n) = \infty$. Suppose that $m(A_n) < \infty$ for all $n \in \mathbb{N}^+$. Write

$$B_1 := A_1,\ B_2 := A_2 \setminus A_1,\ B_3 := A_3 \setminus A_2, \cdots .$$

Then $B_1, B_2, \cdots$, are pairwise disjoint, and

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n,$$

and hence,

$$\begin{aligned} m\big(\lim_{n\to\infty} A_n\big) =& \sum_{n=1}^{\infty} m(B_n) = m(B_1) \\ &+ \sum_{n=2}^{\infty} \big(m(A_n) - m(A_{n-1})\big) = \lim_{n\to\infty} m(A_n) \end{aligned}$$

by the definition of convergence of series.

(b) From the equality

$$A_1 = (A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup \cdots \cup \left(\bigcap_{n=1}^{\infty} A_n \right)$$

and the subtractivity of Lebesgue measure, it follows that

$$\begin{aligned} m(A_1) =& m(A_1 \setminus A_2) + m(A_2 \setminus A_3) + \cdots + m\Big(\bigcap_{n=1}^{\infty} A_n \Big) \\ =& m(A_1) - \lim_{n\to\infty} m(A_n) + m\Big(\bigcap_{n=1}^{\infty} A_n \Big). \end{aligned}$$

Note that $m(A_1) < \infty$, so we have

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

$\square$

Further on, we have the following general properties of Lebesgue measure.

Proposition. Let $\{A_n\} \subset \mathcal{M}$. The following statements hold.

(v) $m(\underline{\lim}_{n\to\infty} A_n) \le \underline{\lim}_{n\to\infty} m(A_n)$.

(vi) If $\sum_{n=1}^{\infty} m(A_n) < \infty$, then $m\big(\overline{\lim}_{n\to\infty} A_n\big) = 0$ and, in particular, $m\big(\lim_{n\to\infty} A_n \big) = 0$ whenever $\{A_n\}$ is convergent.

Proof. (v) From the monotonicity of m, it follows that

$$m\left(\bigcap_{i=k}^{\infty} A_i \right) \le m(A_n), \quad n \ge k.$$

This implies that

$$m\left(\bigcap_{i=k}^{\infty} A_i\right) \le \inf_{n\ge k} m(A_n).$$

By letting $k \to \infty$ in both sides, we obtain that

$$\begin{aligned} m\big(\varliminf_{n\to\infty} A_n\big) &= \lim_{k\to\infty} m\left(\bigcap_{i=k}^{\infty} A_i\right) \\ &\le \lim_{k\to\infty} \inf_{n\ge k} m(A_n) = \varliminf_{n\to\infty} m(A_n). \end{aligned}$$

(vi) Let $\sum_{n=1}^{\infty} m(A_n) < \infty$. From (iv) (b) and the countable subadditivity it follows that

$$\begin{aligned} m\big(\varlimsup_{n\to\infty} A_n\big) =& m\left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} A_n\right) \\ =& \lim_{k\to\infty} m\left(\bigcup_{n=k}^{\infty} A_n\right) \le \lim_{k\to\infty} \sum_{n=k}^{\infty} m(A_n) = 0. \end{aligned}$$

□

Examples. Please give two divergent sequences $\{A_n\}$ and $\{B_n\}$ such that

$$m\big(\varliminf_{n\to\infty} A_n\big) = \varliminf_{n\to\infty} m(A_n)$$

and

$$m\big(\varliminf_{n\to\infty} B_n\big) < \varliminf_{n\to\infty} m(B_n),$$

respectively.

(i) Consider $\{A_n\}_{n=1}^\infty$ given by $A_{2n} := \big[0, 1 - \frac{1}{n}\big]$ and $A_{2n+1} := \big[0, \frac{1}{n}\big]$. It is clear that

$$\varlimsup_{n\to\infty} A_n = [0,1), \quad \varliminf_{n\to\infty} A_n = \{0\}.$$

Thus, $\{A_n\}$ is divergent and

$$m\big(\varliminf_{n\to\infty} A_n\big) = 0 = \varliminf_{n\to\infty} m(A_n).$$

(ii) Consider $\{B_n\}_{n=1}^\infty$ given by $B_{2n} := \big[0, 1 - \frac{1}{n}\big]$ and $B_{2n+1} := [-1, 0]$. It is clear that

$$\varlimsup_{n\to\infty} A_n = [-1,1), \quad \varliminf_{n\to\infty} A_n = \{0\}.$$

Thus, $\{A_n\}$ is divergent and

$$m\big(\varliminf_{n\to\infty} A_n\big) = 0 < 1 = \varliminf_{n\to\infty} m(A_n).$$

Further properties of Lebesgue measurable sets.

Theorem. Let $E \subset \mathbb{R}$. The following statements are equivalent.

(i) E is Lebesgue measurable.

(ii) For every $\epsilon > 0$ there exists an open set $O \supset E$ such that $m^*(O \setminus E) < \epsilon$.

(iii) For every $\epsilon > 0$ there exists a closed set $F \subset E$ such that $m^*(E \setminus F) < \epsilon$.

(iv) For every $\epsilon > 0$ there exist open set O and closed set F with $F \subset E \subset O$ such that $m(O \setminus F) < \epsilon$.

Proof. Note that (iv) is a direct consequence of the statements (ii) and (iii), and the statement (iii) follows from (ii) due to the fact that $O \setminus E^c = E \setminus O^c$. Thus, it suffices to show the equivalence of (i) and (ii).

(i) $\Rightarrow$ (ii) Suppose that E is Lebesgue measurable. Recall that

$$m(E) = m^*(E) = \inf\{m(O) : O \text{ is open and } O \supset E\}.$$

Then by the definition of infimum, for each $\epsilon > 0$, there is an open set $O \supset E$ such that

$$m(O \setminus E) + m(E) = m(O) < m(E) + \epsilon.$$

This yields $m(O \setminus E) < \epsilon$ whenever $m(E) < +\infty$.

If $m(E) = +\infty$, then we turn to consider

$$E_n := E \cap (-n, n), \quad n \in \mathbb{N}^+.$$

Clearly, $m(E_n) < +\infty$ for every $n \in \mathbb{N}^+$. Thus, from above discussion there exists an open set $O_n \supset E_n$ such that

$$m(O_n \setminus E_n) < \epsilon/2^n.$$

Write $O := \bigcup_{n\in\mathbb{N}^+} O_n$. Then O is open and $O \supset E$. In addition,

$$\begin{aligned} O \setminus E =& \left(\bigcup_{n=1}^{\infty} O_n\right) \setminus \left(\bigcup_{m=1}^{\infty} E_m\right) \\ =& \bigcup_{n=1}^{\infty} \left(\bigcap_{m=1}^{\infty} (O_n \setminus E_m)\right) \subset \bigcup_{n=1}^{\infty} (O_n \setminus E_n). \end{aligned}$$

Thus, by the subadditivity of m, we have

$$m^*(O \setminus E) = m(O \setminus E) \leq \sum_{n=1}^{\infty} m(O_n \setminus E_n) \leq \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon.$$

This proves the statement (ii).

(ii) $\Rightarrow$ (i) Suppose that E satisfies the condition (ii). Then, for every $\epsilon = 1/n$ with $n \in \mathbb{N}^+$, there exists an open set $O_n \supset E$ such that

$$m^*(O_n \setminus E) < \epsilon.$$

Write $O := \bigcap_{n\in\mathbb{N}^+} O_n$. Then $O \supset E$ and

$$m^*(O \setminus E) \leq m^*(O_n \setminus E) < 1/n$$

due to the monotonicity of m^*. By letting $n \to \infty$ we obtain $m^*(O \setminus E) = 0$, so that $O \setminus E$ is Lebesgue measurable (WHY). Thus,

$$E = O \setminus (O \setminus E)$$

is Lebesgue measurable as well (WHY). □

Remark. Let $E \subset \mathbb{R}$ with $m(E) < \infty$. From the above discussion, we observe that for every $\epsilon > 0$, there is a finite union $F = \bigcup_{k=1}^{n} I_k$ of closed intervals such that

$$m(E \triangle F) < \epsilon.$$

In the informal formulation of **J. E. Littlewood**,"*every set is nearly a finite union of intervals*".

Remark. The outline of Lebesgue integral:

Open (closed) sets	Lebesgue measurable sets
$\downarrow$	$\downarrow$
Continuous functions	Lebesgue measurable functions
$\downarrow$	$\downarrow$
Riemann integral	Lebesgue integral

§2. Lebesgue measurable functions

Real-valued functions

Definition. Let $E \subset \mathbb{R}$. A function $f : E \to (-\infty, +\infty)$ is called a **(finite) real-valued function** on E (different from the concept of bounded function), while a function $f : E \to [-\infty, +\infty]$ is called an **extended real-valued function** on E.

Lebesgue measurable functions (finite real-valued)

Let f be a finite real-valued function on $\mathbb{R}$. Recall that

- f is continuous $\Leftrightarrow$ $f^{-1}(O) \in \mathcal{O}$ for every open set O.

Proof. Necessity. Suppose that f is continuous. Let $O \subset \mathbb{R}$ be open. Fix $x \in f^{-1}(O)$. Then $f(x) \in O$, and hence, there exists an $\epsilon > 0$ small enough such that $U(f(x), \epsilon) \subset O$. On the other hand, by the continuity of f, there is a $\delta > 0$ such that $f(U(x, \delta)) \subset U(f(x), \epsilon)$. Thus,

$$U(x, \delta) \subset f^{-1}(f(U(x, \delta))) \subset f^{-1}(U(f(x), \epsilon)) \subset f^{-1}(O).$$

This implies that $f^{-1}(O)$ is an open set.

Sufficiency. Suppose that $f^{-1}(O) \in \mathcal{O}$ for every open set O. Fix $x \in \mathbb{R}$ and $\epsilon > 0$. Since $x \in f^{-1}\big(U(f(x), \epsilon)\big) \in \mathcal{O}$, there is a $\delta > 0$ such that

$$U(x, \delta) \subset f^{-1}\big(U(f(x), \epsilon)\big),$$

and hence,

$$f(U(x, \delta)) \subset f\Big(f^{-1}\big(U(f(x), \epsilon)\big)\Big) = U(f(x), \epsilon).$$

This implies that f is continuous at x. Thus, f is continuous on $\mathbb{R}$. $\square$

This motivates us to introduce the following concept of Lebesgue measurable functions.

Definition. (**Lebesgue measurable function**) Let f be a finite real-valued function on E. We call f a **Lebesgue measurable function** on E if $f^{-1}(O) \in \mathcal{M}$ for every open set $O \subset \mathbb{R}$.

Remark. If $f : E \subset \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable on E, then $E \in \mathcal{M}$ (WHY).

Remark. Let $E \subset \mathbb{R}$. Then $E \in \mathcal{M}$ if and only if χ_E is Lebesgue measurable on $\mathbb{R}$. Indeed, it is a direct consequence of the following fact:

$$\mathbb{R}[\chi_E \geq c] = \begin{cases} \emptyset, & 1 < c, \\ E, & 0 < c \leq 1, \\ \mathbb{R}, & c \leq 0. \end{cases}$$

Characteristics of Lebesgue measurable function.

Theorem. Let $f : E \subset \mathbb{R} \to \mathbb{R}$. The following statements are equivalent.

(i) f is Lebesgue measurable on E.

(ii) $f^{-1}(F) \in \mathcal{M}$ for every closed set $F \subset \mathbb{R}$.

Proof. Since O^c is closed for every open set O, and F^c is open for every closed set F, we obtain the conclusion by the definition of Lebesgue measurable functions, immediately. □

Theorem. Let $f : E \subset \mathbb{R} \to \mathbb{R}$, and let $E = E_1 \cup E_2$, where $E_1, E_2 \in \mathcal{M}$ and $E_1 \cap E_2 = \emptyset$. The following statements are equivalent.

(i) f is Lebesgue measurable on E.

(ii) $f|_{E_1}$ and $f|_{E_2}$ are Lebesgue measurable functions on E_1 and E_2, respectively.

Proof. (i) $\Rightarrow$ (ii) Suppose that f is Lebesgue measurable on E. Clearly,

$$E_i[f < a] = E[f < a] \cap E_i, \quad i = 1, 2.$$

This implies that $E_i \in \mathcal{M}$ for $i = 1, 2$ due to the closedness of $\mathcal{M}$ under the operation of intersection.

(ii) $\Rightarrow$ (i) Suppose that f_i is Lebesgue measurable on E_i for $i = 1, 2$. Clearly,

$$E[f < a] = E_1[f < a] \cup E_2[f < a].$$

This implies that $E[f < a] \in \mathcal{M}$ due to the closedness of $\mathcal{M}$ under the operation of union. $\square$

Theorem. Let f be a finite real-valued function on $\mathbb{R}$. The following statements are equivalent.

(i) f is Lebesgue measurable.

(ii) $E[f < a] \in \mathcal{M}$ for each $a \in \mathbb{R}$.

(iii) $E[f > a] \in \mathcal{M}$ for each $a \in \mathbb{R}$.

(iv) $E[f \leq a] \in \mathcal{M}$ for each $a \in \mathbb{R}$.

(v) $E[f \geq a] \in \mathcal{M}$ for each $a \in \mathbb{R}$.

(vi) $E[b < f < a] \in \mathcal{M}$ for each $a, b \in \mathbb{R}$ with $b < a$.

(vii) $E[b \leq f < a] \in \mathcal{M}$ for each $a, b \in \mathbb{R}$ with $b < a$.

(viii) $E[b < f \leq a] \in \mathcal{M}$ for each $a, b \in \mathbb{R}$ with $b < a$.

(ix) $E[b \leq f \leq a] \in \mathcal{M}$ for each $a, b \in \mathbb{R}$ with $b < a$.

Proof. (i) $\Rightarrow$ (ii) Trivial.

(ii) $\Rightarrow$ (i) Suppose that f satisfies the condition (ii). Let $O \in \mathcal{O}$. Write $O := \bigcup_n I_n$ with I_n, construction intervals of O. Since $I_1, I_2, \cdots$ are pairwise disjoint, we have

$$f^{-1}(O) = f^{-1}\left(\bigcup_n I_n\right) = \bigcup_n f^{-1}(I_n).$$

For interval $I_n = (-\infty, a)$ with some $a \in \mathbb{R}$, it is clear that $f^{-1}(I_n) \in \mathcal{M}$.

For interval $I_n = (b, +\infty)$ with some $b \in \mathbb{R}$, since

$$(b, +\infty) = \mathbb{R} \setminus (-\infty, b] = \mathbb{R} \setminus \Big(\bigcap_{n=1}^{\infty}(-\infty, b + 1/n)\Big),$$

we have

$$\begin{aligned} f^{-1}(I_n) =& f^{-1}\Big(\mathbb{R} \setminus \Big(\bigcap_{n=1}^{\infty}(-\infty, b + 1/n)\Big)\Big) \\ =& E \setminus f^{-1}\Big(\bigcap_{n=1}^{\infty}(-\infty, b + 1/n)\Big) \\ =& E \setminus \Big(\bigcap_{n=1}^{\infty} f^{-1}((-\infty, b + 1/n))\Big) \in \mathcal{M}. \end{aligned}$$

For interval $I_n = (a, b)$ with some $a, b \in \mathbb{R}$, we also have $f^{-1}(I_n) \in \mathcal{M}$ due to the following decomposition

$$\begin{aligned} (a, b) =& \mathbb{R} \setminus \big((-\infty, a] \cup [b, +\infty)\big) \\ =& \mathbb{R} \setminus \Big(\big(\bigcap_{n=1}^{\infty}(-\infty, a + 1/n)\big) \cup \big(\bigcap_{n=1}^{\infty}(b - 1/n, \infty)\big)\Big). \end{aligned}$$

Finally, note that

$$\begin{aligned} E[f \leq a] =& \bigcap_{n=1}^{\infty} E[f < a + 1/n], \\ E[f > a] =& E \setminus E[f \leq a], \\ E[f \geq a] =& \bigcap_{n=1}^{\infty} E[f > a - 1/n], \end{aligned}$$

and

$$\begin{aligned} E[b<f<a] =& E[f<a] \setminus E[f \geq b], \\ E[b \leq f<a] =& E[f<a] \setminus E[f > b], \\ E[b<f \leq a] =& E[f \leq a] \setminus E[f \geq b], \\ E[b \leq f \leq a] =& E[f \leq a] \setminus E[f > b]. \end{aligned}$$

Then the equivalence of (ii), (iii), (iv), (v), (vi), (vii), (viii) and (ix) is a direct consequence of the closedness of $\mathcal{M}$ under operations of union, intersection and complement. □

Algebraic properties of Lebesgue measurable function.

Theorem. Let f and g be two Lebesgue measurable functions on $E \subset \mathbb{R}$. Then

(i) αf is Lebesgue measurable for each $\alpha \in \mathbb{R}$.

(ii) $f+g$ is Lebesgue measurable.

(iii) f^2 is Lebesgue measurable.

(iv) fg is Lebesgue measurable.

(v) f/g is Lebesgue measurable whenever $g(x) \neq 0$ for all $x \in E$.

(vi) $\max\{f,g\}$ and $\min\{f,g\}$ are both Lebesgue measurable functions.

Proof. (i) The case for $\alpha = 0$ is trivial. Let $\alpha \neq 0$. It is clear that

$$E[\alpha f < a] = E[f < a/\alpha] \in \mathcal{M},$$

and hence, αf is Lebesgue measurable.

(ii) Note that

$$E[f+g>a] = \bigcup_{r \in \mathbb{Q}} \big(E[f>r] \cap E[r > a-g] \big).$$

This implies that $f+g$ is Lebesgue measurable.

(iii) Note that

$$E[f^2 < a] = \emptyset \in \mathcal{M}$$

whenever $a \leq 0$. Let $a > 0$. Then

$$E[f^2 < a] = E[-a < f < a] \in \mathcal{M}.$$

This proves that f^2 is Lebesgue measurable.

The statement (iv) is a direct consequence of (iii) due to the fact that

$$fg = [(f+g)^2 - (f-g)^2]/4.$$

Since $1/g$ is Lebesgue measurable (WHY), the statement (v) is a direct consequence of (iv).

(vi) Note that

$$E[\max\{f,g\} \geq a] = E[f \geq a] \cup E[g \geq a].$$

Thus, $\max\{f,g\}$ is Lebesgue measurable. In addition, $\min\{f,g\}$ is also Lebesgue measurable due to the fact that

$$\min\{f,g\} = -\max\{-f,-g\}.$$

□

Limit of sequence of Lebesgue measurable functions.

Theorem. Let $\{f_n\}$ be a sequence of Lebesgue measurable function f_n on $E \subset \mathbb{R}$. The following statements hold.

(i) If $\sup_{n\in\mathbb{N}^+} f_n$ is finite, then $\sup_{n\in\mathbb{N}^+} f_n$ is Lebesgue measurable.

(ii) If $\inf_{n\in\mathbb{N}^+} f_n$ is finite, then $\inf_{n\in\mathbb{N}^+} f_n$ is Lebesgue measurable.

(iii) If $\overline{\lim}_{n\to\infty} f_n$ is finite, then $\overline{\lim}_{n\to\infty} f_n$ is Lebesgue measurable.

(iv) If $\underline{\lim}_{n\to\infty} f_n$ is finite, then $\underline{\lim}_{n\to\infty} f_n$ is Lebesgue measurable.

(v) If $\{f_n\}$ converges pointwise to a finite real-valued function f, then f is Lebesgue measurable.

Proof. (i) Note that

$$\begin{aligned}\sup_{n\in\mathbb{N}^+} f_n(x) &= \max_{n\in\mathbb{N}^+} f_n(x)\\ &= \lim_{n\to\infty} \max\{f_1(x), f_2(x), \cdots, f_n(x)\}\\ &= \lim_{n\to\infty} F_n(x),\end{aligned}$$

where $F_n(x) := \max\{f_1(x), f_2(x), \cdots, f_n(x)\}$. Since $\{F_n\}$ is monotone increasing sequence of Lebesgue measurable functions, we have

$$E[\max_{n\in\mathbb{N}^+} f_n > a] = \bigcup_{n=1}^{\infty} E[F_n > a]$$

for each $a \in \mathbb{R}$ (WHY), so that $\max_n f_n$ is Lebesgue measurable.

The statement (ii) follows from (i) and the fact that

$$\min_{n\in\mathbb{N}^+} f_n = -\lim_{n\to\infty} \max\{-f_1(x), -f_2(x), \cdots, -f_n(x)\}.$$

The statement (iii) follows from (i), (ii) and the fact that

$$\overline{\lim_{n\to\infty}} f_n = \inf_{n\in\mathbb{N}^+} \sup_{k\geq n} f_k.$$

The statement (iv) follows from (i), (ii) and the fact that

$$\underline{\lim_{n\to\infty}} f_n = \sup_{n\in\mathbb{N}^+} \inf_{k\geq n} f_k.$$

Finally, the statement (v) is a direct consequence of (iii) or (iv). $\square$

Approximation for Lebesgue measurable functions.

The well-known **Weierstrass approximation theorem** (K. Weierstrass, 1885) states that

• Let $f : [0,1] \to \mathbb{R}$ be continuous. For each $\epsilon > 0$, there is a polynomial p such that $|f(x) - p(x)| < \epsilon$ for all $x \in [0,1]$, i.e.,

$$\|f - p\|_\infty := \sup_{0 \le x \le 1} |f(x) - p(x)| < \epsilon.$$

Compared with the Weierstrass' famous example of 1861 on the existence of a **continuous nowhere differentiable function**, the Weierstrass approximation theorem shows that every continuous function can be approximated by a sequence of polynomials.

What about the Lebesgue measurable function? To end this, we need the following concept of so-called simple functions.

Definition. (**Simple functions**) A function $f : E \subset \mathbb{R} \to \mathbb{R}$ is said to be a **simple function** if f is a linear combinations of characteristic functions of Lebesgue measurable sets, that is to say, there are constants $\alpha_1, \alpha_2, \cdots, \alpha_N$ and Lebesgue measurable sets $E_1, E_2, \cdots, E_N$ such that

$$f = \sum_{n=1}^{N} \alpha_n \chi_{E_n}.$$

Remark. A Lebesgue measurable function f is simple function **if and only if** the range of f is a finite set.

Theorem. (**Approximated by simple functions**) Let $f : E \subset \mathbb{R} \to \mathbb{R}$. If f is Lebesgue measurable, then there exists a sequence $\{f_n\}$ of simple functions on E such that $\{f_n\}$ converges **pointwise** to f on E.

Proof. Suppose that f is Lebesgue measurable. For every $n \in \mathbb{N}^+$ we write

$$E_{n,k} := E[k/n \leq f < (k+1)/n]$$

with $k = -n^2, -n^2+1, \cdots, n^2-1$, and denote

$$f_n(x) := \sum_{k=-n^2}^{n^2-1} \frac{k}{n} \chi_{E_{n,k}}(x), \quad x \in E.$$

It is clear that f_n is simple function since all sets $E_{n,k}$ are Lebesgue measurable.

On the other hand, we can show that such sequence $\{f_n\}$ converges pointwise to f on E. Indeed, let $x \in E$ and take $N \in \mathbb{N}^+$ such that $|f(x)| < N$. Fix natural number $n \geq N$. Then there is a unique natural number $k \in [-n^2, n^2-1]$ such that $x \in E_{n,k}$, so that

$$k/n \leq f(x) < (k+1)/n.$$

In addition, by the definition of f_n we have $f_n(x) = k/n$. Thus,

$$0 \leq f(x) - f_n(x) < 1/n \to 0, \quad n \to \infty.$$

This implies that $f_n(x) \to f(x)$ as $n \to \infty$ for every $x \in E$. $\square$

Further on, if the function f given in the last theorem is bounded on E. Then there is an $N \in \mathbb{N}^+$ such that $\sup_{x \in E} |f(x)| < N$, so that for all $n \geq N$ we have

$$\sup_{x \in E} |f(x) - f_n(x)| < 1/n \to 0, \quad n \to \infty.$$

This implies that $f_n \rightrightarrows f$ as $n \to \infty$. Therefore, we have proved the following result.

Theorem. (**Approximated by simple functions**) Let $f : E \subset \mathbb{R} \to \mathbb{R}$ be **bounded**. If f is Lebesgue measurable, then there exists a sequence $\{f_n\}$ of simple functions on E such that $\{f_n\}$ converges **uniformly** to f on E.

Remark. In the last two approximation theorems, we can choose the desired sequence $\{f_n\}$ of such simple functions f_n, where **each f_n is a linear combination of characteristic functions of Lebesgue measurable sets with finite measure**. It suffices to show that such sequence exists for non-negative measurable function $f : E \to \mathbb{R}$ (WHY). To this end, we write

$$F_n := \begin{cases} f(x), & x \in E \cap I_n \text{ and } f(x) \le n, \\ n, & x \in E \cap I_n \text{ and } f(x) > n, \\ 0, & x \in E \setminus I_n. \end{cases}$$

where $I_n := (-n, n)$ for $n \in \mathbb{N}^+$. Then $f_n \to f$ on E. Further on, write

$$E_{l,j} := \left\{ x \in E \cap I_n : \frac{l}{j} < F_n \le \frac{l+1}{j} \right\}, \quad 0 \le l < jn,$$

and denote

$$F_{n,j}(x) := \sum_{l=0}^{jn-1} \frac{l}{j} \chi_{E_{l,j}}(x), \quad x \in E.$$

Then $m(E_{l,j}) \le m(I_n) = 2n < +\infty$ for each $n \in \mathbb{N}^+$ and

$$0 \le F_n(x) - F_{n,j}(x) \le 1/j, \quad x \in E.$$

Finally, by denote $f_n := F_{n,n}$, we obtain the sequence $\{f_n\}$ desired.

Remark. So far, we have two kinds of convergences of sequences of functions: **pointwise convergence** and **uniform convergence**. There are also some kinds of convergence (for instance, **convergence almost everywhere** and **convergence in measure**) of sequence of functions, which will be discussed in the sequel.

§3. Convergence almost everywhere

In this section, we will introduce some kinds of convergences of sequences of functions and discuss relations of them.

Convergences of sequence $\{f_n\}$

Throughout this section, we suppose that all the functions $f, f_1, f_2, \cdots$ are finite real-valued functions.

1. **Uniform convergence** ($f_n \rightrightarrows f$ on E): for arbitrary $\epsilon > 0$, there is a natural number $N = N(\epsilon)$ such that

$$\sup_{x\in E} |f_n(x) - f(x)| < \epsilon, \quad n \geq N.$$

2. **Pointwise convergence** ($f_n \to f$ on E): for every $x \in E$ fixed and for arbitrary $\epsilon > 0$, there is a natural number $N = N(x, \epsilon)$ such that

$$|f_n(x) - f(x)| < \epsilon, \quad n \geq N.$$

3. **Convergence almost everywhere** ($f_n \xrightarrow{a.e.} f$ on E)

Definition. We say that f_n **converges almost everywhere** (**a.e.**, for short) to f on E if there is a set $E_0 \in \mathcal{M}$ with $m(E_0) = 0$ such that $f_n \to f$ on $E \setminus E_0$.

Remark. $f_n \xrightarrow{a.e.} f$ on E if and only if there is a set $E_0 \in \mathcal{M}$, with $m(E_0) = 0$, satisfies that for every $x \in E \setminus E_0$ fixed and for arbitrary $\epsilon > 0$, there is a natural number $N = N(x, \epsilon)$ such that

$$|f_n(x) - f(x)| < \epsilon, \quad n \geq N.$$

Remark. Let $f_n \xrightarrow{a.e.} f$ on E. If $E_0 \cap E = \emptyset$, then $f_n \to f$ on E.

Theorem. Let $f_1, f_2, \cdots$ be Lebesgue measurable on E. If $f_n \xrightarrow{a.e.} f$ on E, then there is a Lebesgue measurable function h on E such that $f(x) = h(x)$ for a.e. $x \in E$.

Proof. Let $f_n \xrightarrow{a.e.} f$ on E. Then there is a set $E_0 \in \mathcal{M}$, with $m(E_0) = 0$, such that $f_n \to f$ on $E \setminus E_0$. Since $E \setminus E_0 \in \mathcal{M}$, $f|_{E\setminus E_0}$ is a Lebesgue measurable function on $E \setminus E_0$ (WHY). Write

$$h(x) := \begin{cases} f(x), & x \in E \setminus E_0, \\ 0, & x \in E_0. \end{cases}$$

Clearly, h is Lebesgue measurable on E and $f(x) = h(x)$ for a.e. $x \in E$. □

Corollary. Let $f_1, f_2, \cdots$ be Lebesgue measurable on E. If $f_n \xrightarrow{a.e.} f$ on E, then f is Lebesgue measurable on E.

Proof. Note that if $f(x) = g(x)$ for a.e. $x \in E$ and g is Lebesgue measurable on E, then f is Lebesgue measurable on E as well (WHY). The conclusion is a direct consequence of the last theorem. □

Relations of various convergences

It is clear that

$$\text{``}f_n \rightrightarrows f\text{''} \Rightarrow \text{``}f_n \to f\text{''} \Rightarrow \text{``}f_n \xrightarrow{a.e.} f\text{''},$$

however, the inversions are not true. Could you give some counterexamples?

Relation between $f_n \to f$ and $f_n \rightrightarrows f$

Theorem. Let $f_1, f_2, \cdots$ be Lebesgue measurable functions on E with $m(E) < +\infty$. If $f_n \to f$ on E, then for every $\epsilon > 0$, there is a Lebesgue measurable set $\Omega \subset E$ such that $m(E \setminus \Omega) < \epsilon$ and $f_n \rightrightarrows f$ on Ω.

Proof. Let $f_n \to f$ on E, that is, $\lim_{n\to\infty} f_n(x) = f(x)$ for each $x \in E$. Thus, for each $k \in \mathbb{N}^+$, there is an $N \in \mathbb{N}^+$ such that

$$|f_n(x) - f(x)| < 1/k, \quad n \geq N.$$

This motivates us to consider the sets

$$E_{k,N} := \{x \in E : |f_n(x) - f(x)| < 1/k \text{ for all } n \geq N\}$$

with $k, N \in \mathbb{N}^+$.

For each $k \in \mathbb{N}^+$ fixed, note that

$$E_{k,1} \subset E_{k,2} \subset \cdots \subset E_{k,N} \subset E_{k,N+1} \subset \cdots,$$

and hence,

$$\lim_{N\to\infty} E_{k,N} = \bigcup_{N=1}^{\infty} E_{k,N} = E.$$

By the continuity of the Lebesgue measure, we have

$$\lim_{N\to\infty} m(E_{k,N}) = m(E),$$

so that there is an $N_k \in \mathbb{N}^+$ such that

$$m(E \setminus E_{k,N_k}) = m(E) - m(E_{k,N_k}) < 1/2^k$$

due to the fact that $m(E) < +\infty$. Take $K \in \mathbb{N}^+$ such that

$$\sum_{k\geq K} 2^{-k} < \epsilon$$

and write $\Omega := \bigcap_{k\geq K} E_{k,N_k}$. Then $m(E \setminus \Omega) < \epsilon$. Indeed,

$$\begin{aligned} m(E \setminus \Omega) =& m(E \cap \Omega^c) = m\Bigg(\bigcup_{k\geq K} (E \setminus E_{k,N_k})\Bigg) \\ &\leq \sum_{k\geq K} m(E \setminus E_{k,N_k}) \leq \sum_{k\geq K} 2^{-k} < \epsilon. \end{aligned}$$

In addition, it is easy to show that $f_n \rightrightarrows f$ on Ω. Indeed, let $\epsilon' > 0$ and take $k \geq K$ such that $1/k < \epsilon'$. Then for each $x \in E_{k,N_k}$, we have

$$|f_n(x) - f(x)| < 1/k < \epsilon, \quad n \geq N_k.$$

Thus,

$$\sup_{x \in \Omega} |f_n(x) - f(x)| < \epsilon, \quad n \geq N_k.$$

This completes the proof. □

Relation between $f_n \xrightarrow{a.e.} f$ and $f_n \rightrightarrows f$

D. Egorov (1869 - 1931, Russian) give a condition for the uniform convergence of a pointwise convergent sequence of measurable functions as showed in the following

Theorem. (**Egorov**, 1911) Let $f_1, f_2, \cdots$ be Lebesgue measurable functions on E with $m(E) < +\infty$. If $f_n \xrightarrow{a.e.} f$ on E, then for every $\epsilon > 0$, there is a closed subset $F \subset E$ such that $m(E \setminus F) < \epsilon$ and $f_n \rightrightarrows f$ on F.

Proof. Suppose that $f_n \xrightarrow{a.e.} f$ on E. Then there is a zero-measure set $E_0 \in \mathcal{M}$ such that $f_n \to f$ on $E \setminus E_0 := \widetilde{E}$. Let $\epsilon > 0$. Then by the last theorem, there is a Lebesgue measurable set $\Omega \subset \widetilde{E}$ with $m(\widetilde{E} \setminus \Omega) < \epsilon/2$ such that $f_n \rightrightarrows f$ on Ω. Further on, take a closed subset $F \subset \Omega$ such that $m(\Omega \setminus F) < \epsilon/2$. Then $f_n \rightrightarrows f$ on F and

$$\begin{aligned} m(E \setminus F) =& m(\widetilde{E} \setminus F) = m\big((\widetilde{E} \setminus \Omega) \cup (\Omega \setminus F)\big) \\ \leq& m(\widetilde{E} \setminus \Omega) + m(\Omega \setminus F) < \epsilon/2 + \epsilon/2 = \epsilon. \end{aligned}$$

□

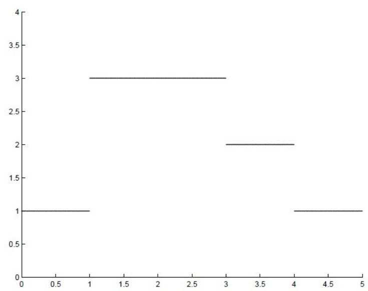
Remark. In the informal formulation of **J. E. Littlewood**, Egorov theorem states that "*every convergent sequence is nearly uniformly convergent*".

§4. Relations between measurable functions and continuous functions

Step functions

Recall that the characteristic function χ_E of a set E is defined by

$$\chi_E(x) := \begin{cases} 1, & x \in E, \\ 0, & x \notin E. \end{cases}$$

For the Riemann integral it is in effect the class of **step functions**. 

Definition. (**Step functions**) A function f is called **step function** if it is of the form

$$f = \sum_{n=1}^{N} \alpha_n \chi_{I_n},$$

where α_n are constants and I_n are (closed) intervals.

Remark. All the step functions are simple functions, however, there is a simple function not step function, for example, $\chi_{\mathbb{Q}}$.

Remark. All the continuous functions, step functions, simple functions as well as monotonic functions are Lebesgue measurable functions.

Proposition. (**Approximated by continuous functions**) Let $f : [a, b] \to \mathbb{R}$ be a step function. Then there is a sequence $\{f_n\}$ of continuous functions on $[a, b]$ such that $f_n \xrightarrow{a.e.} f$ on $[a, b]$.

Proof. It is easy by considering piecewise linear functions. $\square$

Proposition. (**Approximated by step functions**) Let $f : E \to \mathbb{R}$ be a simple function. Then there is a sequence $\{f_n\}$ of step functions such that $f_n \xrightarrow{a.e.} f$ on E.

Proof. It suffices to show that for each Lebesgue measurable subset A with $m(A) < +\infty$, the function χ_A can be approximated a.e. by step functions (WHY). To this end, let $\epsilon > 0$ and take finite union $\bigcup_{n=1}^N I_n$ of disjoint open intervals such that

$$m\left(A \triangle \bigcup_{n=1}^{N} I_n\right) < \epsilon.$$

It is clear that

$$\chi_A(x) = \sum_{n=1}^{N} \chi_{I_n}(x)$$

except possibly on a set of measure less than ϵ. Thus, for special value $\epsilon = 1/k$, there is a step function f_k such that $m(E_k) < 1/2^k$, where

$$E_k := \{x : f(x) \neq f_k(x)\}.$$

Now write

$$F := \bigcap_{k=1}^{\infty} \bigcup_{n=k+1}^{\infty} E_n.$$

Then for $k \in \mathbb{N}^+$, we have

$$\begin{aligned} m(F) \leq & m\left(\bigcup_{n=k+1}^{\infty} E_n\right) \leq \sum_{n=k+1}^{\infty} m(E_n) \\ & \leq \sum_{n=k+1}^{\infty} 1/2^n = 1/2^k \to 0, \quad k \to \infty. \end{aligned}$$

This implies that $m(F) = 0$.

On the other hand, note that

$$F^c = \bigcup_{k=1}^{\infty} \bigcap_{n=k+1}^{\infty} E_n^c,$$

that is, for each $x \in F^c$, there is a $K \in \mathbb{N}^+$ such that $f(x) = f_n(x)$ for all $n \geq K+1$. Thus, $f_n(x) \to f(x)$ as $n \to \infty$, and hence, $f_n \xrightarrow{a.e.} f$ on E. $\square$

By the last proposition, we obtain the following approximation theorem for Lebesgue measurable functions, immediately.

Theorem. (**Approximated by step functions**) Let $f : E \to \mathbb{R}$ be Lebesgue measurable on E. Then there exists a sequence $\{f_n\}$ of step functions such that $f_n \xrightarrow{a.e.} f$ on E.

Relations between measurable functions and continuous functions

Recall that for a finite real-valued function f continuous on open interval $I = (a, b)$, we can take a closed interval $F \subset I$ such that $l(I \backslash F) < \epsilon$ and $f|_F$ is uniformly continuous on F. For example, take $F = [a + \epsilon/3, b - \epsilon/3]$.

For Lebesgue measurable functions, we have the following result due to **Nikolai Lusin** (or **Luzin**) (1883 - 1950, Russian).

Theorem. (**Lusin**, 1912) Let $f : E \subset \mathbb{R} \to \mathbb{R}$ with $m(E) < +\infty$. If f is Lebesgue measurable on E, then for each $\epsilon > 0$ there exists a closed subset $F \subset E$ with $m(E \setminus F) < \epsilon$ such that $f|_F$ is continuous on F.

Proof. Suppose that f is Lebesgue measurable on E. Take a sequence $\{f_n\}$ of step functions such that $f_n \xrightarrow{a.e.} f$ on E. Then for each $n \in \mathbb{N}^+$, there is a set E_n such that

$$m(E_n) < 1/2^n$$

and that f_n is continuous outside E_n. By Egorov theorem, there is a set F_ϵ with $m(E \setminus F_\epsilon) < \epsilon/3$ and $f_n \rightrightarrows f$ on F_ϵ. Now take $N \in \mathbb{N}^+$ such that $\sum_{n \geq N} 1/2^n < \epsilon/3$ and write

$$\widetilde{F} := F_\epsilon \setminus \left(\bigcup_{n \geq N} E_n \right).$$

Note that f_n is continuous on $\widetilde{F}$ for all $n \geq N$, so that f is continuous on $\widetilde{F}$ as well (WHY). Finally, take a closed subset $F \subset \widetilde{F}$ such that $m(\widetilde{F} \setminus F) < \epsilon/3$. Then

$$\begin{aligned} m(E \setminus F) =& m(E \setminus \widetilde{F}) + m(\widetilde{F} \setminus F) \\ \leq& m(E \setminus F_\epsilon) + m\left(\bigcup_{n \geq N} E_n \right) + m(\widetilde{F} \setminus F) \\ \leq& m(E \setminus F_\epsilon) + \sum_{n \geq N} m(E_n) + m(\widetilde{F} \setminus F) \\ <& \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon. \end{aligned}$$

□

Remark. In the informal formulation of **J. E. Littlewood**, Lusin theorem states that "*every function is nearly continuous*".

Summary. Although the notations of Lebesgue measurable sets and measurable functions represent new tools, we should not overlook their relations to the older concepts: **intervals** and **continuous functions**, which are essential in the theory of **Riemann integrals**. The following Littlewood three principles provide a useful intuitive guide in the initial study of the theory of Lebesgue measurable sets and measurable functions:

(i) *Every measurable set is nearly a finite union of intervals.*

(ii) *Every measurable function is nearly continuous.*

(iii) *Every convergent sequence of measurable functions is nearly uniformly convergent.*

In the next chapter, we will travel on the core theory of real analysis - Lebesgue integrals.

Chapter 3: Lebesgue integrals

§1. Some simple cases

Lebesgue integrals of bounded Lebesgue measurable functions on $[a, b]$.

In this subsection, we always consider the bounded closed interval $E := [a, b]$ and construct a "new" kind of integrals for Lebesgue measurable functions on E.

Note that for the function f bounded on E, there are constants m, M such that

$$m < f(x) < M, \quad x \in E.$$

With an analogue of the Riemann integral, we can define the Lebesgue integral of f as follows.

Definition. Let f be a bounded Lebesgue measurable function on E. For every partition $D : m = y_0 < y_1 < \cdots < y_{n-1} < y_n = M$ of $[m, M)$, denote

$$\delta(D) := \max_{1 \leq k \leq n} (y_k - y_{k-1}), \quad E_k := E[y_{k-1} \leq f < y_k].$$

Further on, take $\eta_k \in [y_{k-1}, y_k]$ arbitrary and write

$$S := \sum_{k=1}^{n} \eta_k m(E_k).$$

If the limit

$$\lim_{\delta(D) \to 0} S := s$$

exists and s is independent of the partition D and the points η_k, then we call f is **Lebesgue integrable** on E. In addition, we call s the **Lebesgue integral** of f on the set E

and write

$$s := \int_E f(x)\,dx = (L)\int_a^b f(x)\,dx.$$

Remarks. (i) The Riemann integral of f on $[a, b]$ (**if exists**) is denoted by

$$(R)\int_a^b f(x)\,dx.$$

(ii) Clearly, the "**If** part" in the definition above is equivalent to that "there is a constant s satisfies that for any $\epsilon > 0$, there is a $\delta > 0$ such that, for any partition D with $\delta(D) < \delta$,

$$|S(D) - s| = \left|\sum_{k=1}^{n} \eta_k m(E_k) - s\right| < \epsilon$$

for all $\eta_k \in [y_{k-1}, y_k]$.

(iii) Let f be a bounded Lebesgue measurable function on E. By the Heine theorem, f is Lebesgue integrable on E if and only if the limit $\lim_{n\to\infty} S(D_n)$ exists for every sequence $\{D_n\}$ of partitions of $[m, M)$ with $\delta(D_n) \to 0$ as $n \to \infty$ and is independent of D and η_k.

Uniqueness of the Lebesgue integral

Theorem. Let f be Lebesgue integral on E. Then the Lebesgue integral $\int_E f(x)\,dx$ is independent of the choice of m and M.

Proof. Write $a := \inf_{x\in E} f(x)$ and $b := \sup_{x\in E} f(x)$. It suffices to consider the case

$$m < a \le f(x) \le b < M.$$

Let D be a partition of $[m, M)$:

$$m = y_0 < y_1 < \cdots < y_{n-1} < y_n = M.$$

Then there are k_0 and k_1 such that $a \in [y_{k_0-1}, y_{k_0})$ and $b \in [y_{k_1-1}, y_{k_1})$, so that

$$S(D) = \sum_{k=1}^{n} \eta_k m(E_k) = \sum_{k=k_0}^{k_1} \eta_k m(E_k)$$

since such sets E_k are all empty sets for $k < k_0$ and $k > k_1$. Further on, note that

$$S(D) = \sum_{k=k_0}^{k_1} \eta_k m(E_k) = \sum_{k=k_0}^{k_1-1} \eta_k m(E_k) + bm(E')$$

where $E' := E[f = b]$. By taking $y_{k_0} = a$ and $y_{k_1} = b$, we obtain a partition D' of $[a, b)$:

$$a = y_{k_0} < y_1 < \cdots < y_{n-1} < y_{k_1} = b.$$

and

$$S(D) = S(D') + bm(E').$$

Thus, the integral of f is dependent only if the value a and b, so that independent of m and M. This completes the proof. $\square$

Examples.

(i) $(L) \int_a^b dx = (R) \int_a^b dx = b - a$.

(ii) The Dirichlet function D is not Riemann integrable on $[0, 1]$, while Lebesgue integrable on $[0, 1]$ with

$$(L) \int_0^1 D(x)\, dx = 0.$$

(iii) Let

$$f(x) = \begin{cases} 1, & 0 \le x \le \frac{1}{2}; \\ 2, & \frac{1}{2} \le x \le 1. \end{cases}$$

Then f is Lebesgue integrable on $[0,1]$ and

$$(L)\int_0^1 f(x)\,dx = (R)\int_0^1 f(x)\,dx = 3/2.$$

Proof. (iii) For each partition

$$D : 1 = y_0 < y_1 < \cdots < y_{n-1} < y_n = 3$$

of $[1,3)$, we have

$$\underline{1 = y_0 < y_1} < \cdots < \underline{y_{k_0-1} \leq 2 < y_{k_0}} < \cdots < y_n$$

for some $k_0 \in \{1, 2, \cdots, n\}$. Notice that

$$E_k = \emptyset, \quad k \neq 1, k_0,$$

while $E_1 = [0, 1/2)$ and $E_{k_0} = [1/2, 1]$, so that

$$S(D) = \sum_{k=1}^{n} \eta_k m(E_k) = \eta_1 m(E_1) + \eta_{k_0} m(E_{k_0}) = \frac{\eta_1 + \eta_{k_0}}{2}.$$

Obviously, $\eta_1 \to 1$ and $\eta_{k_0} \to 2$ as $\delta(D) \to 0$. Thus,

$$(L)\int_0^1 f(x)\,dx = \lim_{\delta(D)\to 0} S(D) = 3/2.$$

$\square$

Remark. It is necessary to point out that there are some alternative approaches to define Lebesgue integral, for instance, by the method of approximation oriented from simple functions, see Stein "Real Analysis, Chapter 2".

As showed in precious example (ii), there is a bounded measurable function which is not Riemann integrable, however, every bounded measurable function on E is necessarily Lebesgue integrable.

Theorem. Let f be a Lebesgue measurable function on E. If f is bounded on E, then f is Lebesgue integrable on E.

Proof. By definition, we consider the existence of the limit of sum

$$S(D) = \sum_{k=1}^{n} \eta_k m(E_k)$$

for every partition $D : m = y_0 < y_1 < \cdots < y_{n-1} < y_n = M$. Noticing that the sum $S(D)$ is dependent on the partition D as well as points η_k, we turn to consider two sums $\underline{S}(D), \overline{S}(D)$ which are dependent only in the partition D:

$$\underline{S}(D) := \sum_{k=1}^{n} y_{k-1} m(E_k), \quad \overline{S}(D) := \sum_{k=1}^{n} y_k m(E_k).$$

Note that $m(E) < \infty$ and f is bounded on E, sets

$$A := \{\underline{S}(D) : D \text{ is a partition of } [m, M]\},$$
$$B := \{\overline{S}(D) : D \text{ is a partition of } [m, M]\}$$

are both bounded. Thus,

$$\underline{S} := \sup_D \underline{S}(D) < \infty, \quad \overline{S} := \inf_D \overline{S}(D) < \infty.$$

Now we complete the proof by the following four steps:

(i) $\underline{S}(D_1) \le \underline{S}(D_2)$ and $\overline{S}(D_1) \ge \overline{S}(D_2)$ for $D_1 \subset D_2$;

(ii) $\underline{S}(D_1) \le \overline{S}(D_2)$ for all partitions D_1 and D_2;

(iii) $\underline{S} = \overline{S} := s$;

(iv) s is the Lebesgue integral value of f on E. □

The following estimates of the Lebesbue integral are of importance in the sequel.

Lemma. Let f be a Lebesgue measurable function on E. If $m \le f \le M$ then

$$m \cdot m(E) \le \int_E f(x)\, dx \le M \cdot m(E)$$

Proof. Let $\epsilon > 0$. Clearly, $f(E) \subset [m, M+\epsilon)$, and for each partition D of $[m, M+\epsilon)$: $m = y_0 < y_1 < \cdots < y_{n-1} < y_n = M + \epsilon$,

$$m \cdot m(E) \leq \sum_{k=1}^{n} \eta_k m(E_k) \leq (M+\epsilon) \cdot m(E).$$

By taking limit $\delta(D) \to 0$ we obtain

$$m \cdot m(E) \leq \int_E f(x)\,dx \leq (M+\epsilon) \cdot m(E),$$

then we obtain the conclusion by letting $\epsilon \to 0$. □

We are familiar with the following properties of integrals in the sense of Riemann.

Theorem. (Finite additivity) Let f be a bounded Lebesgue measurable function on E. Then

$$\int_E f(x)\,dx = \sum_{k=1}^{n} \int_{F_k} f(x)\,dx,$$

where $F_1, F_2, \cdots, F_n$ are Lebesgue measurable subsets of E, $F_i \cap F_j = \emptyset$ $(i \neq j)$ and $E = \bigcup_{k=1}^n F_k$.

Proof. It sufficient to prove the conclusion for $n = 2$. Let $n = 2$. Clearly, f is Lebesgue integrable on F_1, F_2. Let D be a partition of E: $m = y_0 < y_1 < \cdots < y_{n-1} < y_n = M$. Denote

$$E_k := E(y_{k-1} \leq f < y_k)$$

and

$$E_{k1} := F_1(y_{k-1} \leq f < y_k),$$
$$E_{k2} := F_2(y_{k-1} \leq f < y_k).$$

Then $E_{k1} \cap E_{k2} = \emptyset$ and $E_k = E_{k1} \cup E_{k2}$. Notice that

$$\sum_{k=1}^{n} \eta_k m(E_k) = \sum_{k=1}^{n} \eta_k m(E_{k1}) + \sum_{k=1}^{n} \eta_k m(E_{k2})$$

By letting $\delta(D) \to 0$ we obtain
$$\int_E f(x)\,dx = \int_{F_1} f(x)\,dx + \int_{F_2} f(x)\,dx.$$
□

Corollary. Let f be Lebesgue integrable on E and let $E_0 \subset E$. If $m(E_0) = 0$, then f is Lebesgue integrable on $E \setminus E_0$ and
$$\int_E f(x)\,dx = \int_{E \setminus E_0} f(x)\,dx.$$

Proof. This is a direct consequence of the theorem above once we observe that
$$\int_{E_0} f(x)\,dx = 0$$
by the definition of integrals. □

Remark. This corollary states that the values of f on a zero-measure subset $E_1 \subset E$ does not impact the integral of f on E. From this point of view, we can observe that the Dirichlet function is integrable on $[0,1]$, immediately.

Example. (iii') Let f be a step function on E, i.e.,
$$f(x) = \begin{cases} c_1, & a \le x \le a_1; \\ c_2, & a_1 < x \le a_2; \\ \cdots \\ c_n, & a_{n-1} < x \le b \end{cases}$$
for some constants $c_1, c_2, \cdots, c_n$. Then f is Lebesgue integrable and by the finite additivity of Lebesgue integral we have
$$(L)\int_a^b f(x)\,dx = \sum_{k=1}^{n} c_k(a_k - a_{k-1}) = (R)\int_a^b f(x)\,dx.$$

Theorem. (Linear property) Let f, g be two bounded Lebesgue measurable functions on E. Then

$$\int_E \Big(af(x) + bg(x)\Big)\,dx = a\int_E f(x)\,dx + b\int_E g(x)\,dx$$

for all $a, b \in \mathbb{R}$.

Proof. It sufficient to prove that

$$\int_E af(x) = a\int_E f(x)\,dx$$

and

$$\int_E \Big(f(x) + g(x)\Big)\,dx = \int_E f(x)\,dx + \int_E g(x)\,dx$$

The first equality follows from the definition of the Lebesgue integral, immediately. For the second equality, we take m, M such that $m \le f, g, f+g < M$ and consider each partition D: $m = y_0 < y_1 < \cdots < y_{n-1} < y_n = M$. Denote

$$E_{ij} := E(y_{i-1} \le f < y_i, y_{j-1} \le g < y_j)$$

for $i, j = 1, \cdots, n$. Then E_{ij}s are disjoint and $E = \bigcup_{i,j=1}^n E_{ij}$. Thus, by the Lemma above we have

$$\begin{aligned}
&\int_{E_{ij}} (f(x) + g(x))\,dx \le (y_i + y_j)m(E_{ij}) \\
=&(y_i - y_{i-1} + y_j - y_{j-1} + y_{i-1} + y_{j-1})m(E_{ij}) \\
\le&(2\delta(D) + y_{i-1} + y_{j-1})m(E_{ij}) \\
=&2\delta(D)m(E_{ij}) + \int_{E_{ij}} y_{i-1}\,dx + \int_{E_{ij}} y_{j-1}\,dx \\
\le&2\delta(D)m(E_{ij}) + \int_{E_{ij}} f(x)\,dx + \int_{E_{ij}} g(x)\,dx
\end{aligned}$$

for all $i, j = 1, \cdots, n$. Therefore,

$$\int_E (f(x) + g(x))\,dx \le 2\delta(D)m(E) + \int_E f(x)\,dx + \int_E g(x)\,dx.$$

By letting $\delta(D) \to 0$ we obtain

$$\int_E (f(x) + g(x))\,dx \leq \int_E f(x)\,dx + \int_E g(x)\,dx.$$

The inverse inequality can also be obtained by analogous argument. Indeed,

$$\begin{aligned}
&\int_{E_{ij}} f(x)\,dx + \int_{E_{ij}} g(x)\,dx \leq (y_i + y_j)m(E_{ij}) \\
=&(y_i - y_{i-1} + y_j - y_{j-1})m(E_{ij}) + (y_{i-1} + y_{j-1})m(E_{ij}) \\
\leq&2\delta(D)m(E_{ij}) + \int_{E_{ij}} (y_{i-1} + y_{j-1})\,dx \\
\leq&2\delta(D)m(E_{ij}) + \int_{E_{ij}} (f(x) + g(x))\,dx,
\end{aligned}$$

so that the finite additivity gives

$$\begin{aligned}
&\int_E f(x)\,dx + \int_E g(x)\,dx \\
=&\sum_{i,j=1}^{n} \left(\int_{E_{ij}} f(x)\,dx + \int_{E_{ij}} g(x)\,dx \right) \\
\leq&2\delta(D) \sum_{i,j=1}^{n} m(E_{ij}) + \sum_{i,j=1}^{n} \int_{E_{ij}} (f(x) + g(x))\,dx \\
=&2\delta(D)m(E) + \int_E (f(x) + g(x))\,dx.
\end{aligned}$$

By letting $\delta(D) \to 0$ we obtain the inequality desired. $\square$

In the sequel, we give some estimates of the Lebesgue integrals.

Lemma. Let f be a bounded Lebesgue measurable function on E and let $f \geq 0$ (a.e.). Then

$$\int_E f(x)\,dx \geq 0.$$

In particular, if $\int_E f(x)\,dx = 0$ then $f = 0$ (a.e.).

Proof. Denote

$$E_1 := E(f \geq 0), \quad E_2 := E(f < 0).$$

It is clear that $E = E_1 \cup E_2$, $E_1 \cap E_2 = \emptyset$ and $m(E_2) = 0$. By the finite additivity and the lemma above, we have

$$\int_E f(x)\,dx = \int_{E_1} f(x)\,dx + \int_{E_2} f(x) = \int_{E_1} f(x)\,dx \geq 0.$$

For the second statement, let $\int_E f(x)\,dx = 0$. It sufficient to prove $m(E(f > 0)) = 0$. Notice that

$$E(f > 0) = \bigcup_{n=1}^{\infty} E(f \geq 1/n),$$

then it suffices to prove $m(E(f \geq 1/n)) = 0$ for each $n \in N$. Indeed, by the finite additivity and the first lemma of this section, we have

$$\begin{aligned} 0 = \int_E f(x)\,dx &= \int_{E(f \geq 1/n)} f(x)\,dx + \int_{E(f<1/n)} f(x)\,dx \\ &\geq \int_{E(f \geq 1/n)} f(x)\,dx \geq m(E(f \geq 1/n))/n, \end{aligned}$$

so that $m(E(f \geq 1/n)) = 0$ for each $n \in N$. □

Remark. For a non-negative bounded function f on E, the lemma above reads that $f = 0$ (a.e.) on E if and only if $\int_E f(x)\,dx = 0$.

Theorem. (Monotonicity) Let f, g be two bounded Lebesgue measurable functions on E. If $f \leq g$ (a.e.), then

$$\int_E f(x)\,dx \leq \int_E g(x)\,dx.$$

In particular,

$$\int_E f(x)\,dx = \int_E g(x)\,dx.$$

whenever $f = g$ (a.e.).

Proof. It is a direct consequence of the lemma above by considering $h := g - f$. $\square$

Corollary. Let f be a bounded Lebesgue measurable function on E. Then

$$\left|\int_E f(x)\,dx\right| \leq \int_E |f(x)|\,dx.$$

Proof. The conclusion is clear once we observe that

$$-|f| \leq f \leq |f|.$$

$\square$

Relations between Lebesgue integral and Riemann integral.

Theorem. Let f be a function defined on $[a,b]$. If f is Riemann integrable then f is Lebesgue integrable; moreover,

$$(L)\int_a^b f(x)\,dx = (R)\int_a^b f(x)\,dx. \tag{0.5}$$

Outline of the proof:

(i) f is bounded;

(ii) f is Lebesgue measurable;

(iii) The equality (0.5) holds.

Proof. (i) We prove it by contradiction. Suppose that f is unbounded on E. Write

$$(R)\int_a^b f(x)\,dx := A$$

and let $\epsilon = 1$. By the definition of Riemann integral, there is a $\delta > 0$ such that for any partition $D : a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ with $\delta(D) := \max_{1\le k\le n}(x_k - x_{k-1}) < \delta$ and points $\xi_k \in [x_{k-1}, x_k]$, we have

$$\left|\sum_{k=1}^n f(\xi_k)(x_k - x_{k-1}) - A\right| < 1,$$

so that

$$\left|\sum_{k=1}^n f(\xi_k)(x_k - x_{k-1})\right| < |A| + 1,$$

Notice that f is unbounded on $[x_{k-1}, x_k]$ for some k due to our hypothesis. For $\xi_1, \cdots, \xi_{k-1}, \xi_{k+1}, \cdots, \xi_n$ fixed, we can choose ξ_k such that

$$\left| \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) \right| > |A| + 1.$$

This deduces a contradiction. Thus, f is necessarily bounded on E.

(ii) Since f is Riemann integral, we can choose a sequence $\{D_k\}$ of partition of E with $D_k \subset D_{k+1}$ and $\delta(D_k) \to 0$ as $k \to \infty$, where

$$D_k : a = x_0^{(k)} < x_1^{(k)} < \cdots < x_{n_k-1}^{(k)} < x_{n_k}^{(k)} = b.$$

Denote

$$m_j^{(k)} := \inf_{x_{j-1}^{(k)} \le x \le x_j^{(k)}} f(x), \quad M_j^{(k)} := \sup_{x_{j-1}^{(k)} \le x \le x_j^{(k)}} f(x),$$

and write

$$\varphi_k(x) = \begin{cases} m_j^{(k)}, & x_{j-1}^{(k)} < x \le x_j^{(k)}; \\ f(a) = a, & x = a. \end{cases}$$

and

$$\psi_k(x) = \begin{cases} M_j^{(k)}, & x_{j-1}^{(k)} < x \le x_j^{(k)}; \\ f(a) = a, & x = a. \end{cases}$$

It is clear that φ_k and ψ_k are all measurable functions and

$$\varphi_1 \le \cdots \le \varphi_k \le \cdots \le f \le \cdots \le \psi_k \le \cdots \le \psi_1,$$

so that $\underline{f} := \lim_{k\to\infty} \varphi_k$ and $\overline{f} := \lim_{k\to\infty} \psi_k$ are both measurable functions and $\underline{f} \le f \le \overline{f}$. In addition, note that

$$\sum_{j=1}^{n} m_j^{(k)}(x_j^{(k)} - x_{j-1}^{(k)}) = (L)\int_a^b \varphi_k(x)\,dx \le (L)\int_a^b \underline{f}(x)\,dx$$

$$\le (L)\int_a^b \overline{f}(x)\,dx \le (L)\int_a^b \psi_k(x)\,dx = \sum_{j=1}^{n} M_j^{(k)}(x_j^{(k)} - x_{j-1}^{(k)}),$$

and notice that the terms in the end-sides are Darboux sums, by letting $k \to \infty$ we obtain

$$\begin{aligned}(L)\int_a^b \underline{f}(x)\,dx =&(L)\int_a^b \overline{f}(x)\,dx \\ =&(R)\int_a^b f(x)\,dx.\end{aligned} \tag{0.6}$$

Therefore,

$$\underline{f} = f = \overline{f} \quad \text{(a.e.)} \tag{0.7}$$

which implies that f is a measurable function since $(\mathbb{R}, \mathbf{L}, m)$ is a complete measure space.

(iii) By (i) and (ii) we can see that f is Lebesgue integrable. Obviously, (0.5) is a direct consequence of (0.6) and (0.7). $\square$

Lebesgue integrals of bounded Lebesgue measurable functions on $E \subset \mathbb{R}$ with $m(E) < +\infty$.

Let $E \in \mathcal{M}$ with $\mu(E) < +\infty$, and let f be a bounded Lebesgue measurable function on E. Then there are constants m, M such that $m \leq f(x) < M$ for all $x \in E$. We can also introduce the Lebesgue integral for such functions following the procedure given in previous subsection.

Definition. For every partition $D : m := y_0 < y_1 < \cdots < y_{n-1} < y_n := M$, denote

$$\delta(D) := \max_{1 \leq k \leq n} (y_k - y_{k-1}), \quad E_k := E(y_{k-1} \leq f < y_k).$$

Further on, take $\eta_k \in [y_{k-1}, y_k]$ arbitrary and write

$$S := \sum_{k=1}^{n} \eta_k m(E_k).$$

If the limit

$$\lim_{\delta(D) \to 0} S := s$$

exists and the value s is independent of D and η_k, then we call f is **Lebesgue integrable** on E. In addition, we call s the **Lebesgue integral** of f on E, and write

$$s := \int_E f(x)\, dx.$$

Remark. All the results in last subsection hold for this kind of integrals. The proofs are left to the reader as exercises.

Examples. (iv) Let E be a Lebesgue measurable set with $\mu(E) < +\infty$. Then

$$\int_E dx = m(E).$$

Proof. Consider the interval $[m, M) = [1, 2)$. For each partition $D : 1 = y_0 < y_1 < \cdots < y_n = 2$, the sum

$$\sum_{k=1}^{n} \eta_k m(E_k) = \eta_1 m(E),$$

where $\eta_1 \in [1, y_1)$. Clearly, $\eta_1 \to 1$ as $\delta(D) \to 0$, so that

$$\int_E dx = \lim_{\delta(D)\to 0} \sum_{k=1}^{n} \eta_k m(E_k) = m(E).$$

$\square$

Remark. This example shows that the integral of a constant function $f = c$ on E equals $c \cdot m(E)$.

(v) Let f be a Lebesgue measurable function on E with $m(E) = 0$. Then

$$\int_E f(x)\, dx = 0.$$

§2. Integrals of measurable functions on measurable subsets of the line

In this section, we consider general Lebesgue measurable subset $E \subset \mathbb{R}$ (**not** necessarily $m(E) < +\infty$).

Truncated functions.

Definition. Let f be a non-negative function on E. For $N > 0$ arbitrary, we define a **truncated function** $[f]_N$ of f by

$$[f]_N(x) := \begin{cases} f(x), & f(x) \leq N; \\ N, & f(x) > N, \end{cases}$$

or equivalently, $[f]_N = \min\{f, N\}$.

Properties. Let f and g be two non-negative functions on E. The following statements hold.

(i) $f_{N_1} \leq f_{N_2} \leq f$ for all $0 < N_1 \leq N_2 < \infty$.

(ii) $[f+g]_N \leq [f]_N + [g]_N \leq [f+g]_{2N}$ for all $N > 0$.

(iii) $[af]_N = a[f]_{N/a}$ for all $a > 0$.

Proof. (i) Let $0 < N_1 \leq N_2$. Notice that

$$[f]_{N_2}(x) := \begin{cases} f(x), & f(x) \leq N_1; \\ f(x), & N_1 < f(x) \leq N_2; \\ N_2, & f(x) > N_2. \end{cases}$$

The conclusion is obvious.

(ii) Let $N > 0$ and $x \in E$. If $f(x), g(x) \leq N$, then

$$\begin{aligned} [f+g]_N(x) \leq& (f+g)(x) \\ =& f(x) + g(x) = [f]_N(x) + [g]_N(x); \end{aligned}$$

otherwise, if $f(x) > N$ (or $g(x) > N$), then $(f+g)(x) > N$, so that

$$[f+g]_N(x) = N \leq N + g(x)(\text{ or } f(x)) \leq f(x) + g(x).$$

This proves the first inequality. The second inequality is obvious due to the facts that $[f]_N = \min\{f, N\}$ and $[g]_N = \min\{g, N\}$. Indeed,

$$[f]_N + [g]_N \leq \min\{f+g, 2N\} = [f+g]_{2N}.$$

$\square$

Remarks. (a) The statement (i) shows that $[f]_N$ is increasing with respect to the index $N > 0$.

(b) The first inequality in (ii) reads that the truncated operator $[\cdot]_N$ is sublinear, i.e., $[f+g]_N \leq [f]_N + [g]_N$; while the second inequality states that $[f]_N + [g]_N$ is **not too large**, actually, it can be dominated by $[f+g]_{2N}$.

Integrals for non-negative functions.

Notice that we can always choose a monotonic increasing sequence $\{E_n\}$ satisfying $E = \bigcup_{n=1}^{\infty} E_n$ and $m(E_n) < \infty$ for all $n \in \mathbb{N}$. For example, take $E_n = E \cap [-n, n]$. We call such sequence $\{E_n\}$ **the finite monotonic cover of the measure of** E.

Remark. The finite monotonic cover of $m(E)$ is **not** unique.

Lemma. Let $E \subset \mathbb{R}$ be a measurable set, f a non-negative measurable function on E, $\{E_n^{(1)}\}$ and $\{E_n^{(2)}\}$ two finite monotonic covers of the measure of E and let $\{M_n^{(1)}\}$ and $\{M_n^{(2)}\}$ be two sequences of positive numbers with $\lim_{n\to\infty} M_n^{(i)} = \infty$ $(i = 1, 2)$. If the limit

$$s_1 := \lim_{n\to\infty} \int_{E_n^{(1)}} [f]_{M_n^{(1)}}(x)\, dx < \infty,$$

then

$$\lim_{n\to\infty} \int_{E_n^{(2)}} [f]_{M_n^{(2)}}(x)\, dx = s_1.$$

Proof. Fix $k \in \mathbb{N}$. Then there is $n_k \in \mathbb{N}$ such that $M_k^{(2)} \leq M_n^{(1)}$ for all $n \geq n_k$, since $\lim_{n\to\infty} M_n^{(1)} = \infty$. Thus, by the monotonicity of the integral, we have

$$\begin{aligned}
\int_{E_k^{(2)}} [f]_{M_k^{(2)}}(x)\, dx &= \int_{E_k^{(2)} \cap E_n^{(1)}} [f]_{M_k^{(2)}}(x)\, dx \\
&\quad + \int_{E_k^{(2)} \setminus E_n^{(1)}} [f]_{M_k^{(2)}}(x)\, dx \\
&\leq \int_{E_n^{(1)}} [f]_{M_n^{(1)}}(x)\, dx \\
&\quad + M_k^{(2)} \cdot m(E_k^{(2)} \setminus E_n^{(1)}) \to s_1
\end{aligned}$$

as $n \to \infty$, since $\lim_{n\to\infty} m(E_k^{(2)} \setminus E_n^{(1)}) = 0$. Thus,

$$s_2 := \lim_{k\to\infty} \int_{E_k^{(2)}} [f]_{M_k^{(2)}}(x)\, dx \le s_1.$$

On the other hand, by the symmetry of above program we have $s_1 \le s_2$. This completes the proof. $\square$

We are ready to give the definition of integral of f on E by the approach of approximation, which is an analogue of that for the improper Riemann integral.

Definition. Let $E \subset \mathbb{R}$ be a Lebesgue measurable set, f a non-negative Lebesgue measurable function on E, $\{E_n\}$ a finite monotonic cover of the measure of E, and let $\{M_n\}$ be a sequence of positive numbers with $\lim_{n\to\infty} M_n = \infty$. We call f is **Lebesgue integrable** on E if the limit

$$s := \lim_{n\to\infty} \int_{E_n} [f]_{M_n}(x)\, dx < \infty,$$

and call s the **Lebesgue integral** of f on E which is denoted by

$$s = \int_E f(x)\, dx.$$

Remarks. (i) Obviously, the integral of f on E defined above is unique, that is, s is independent of choices of $\{E_n\}$ and $\{M_n\}$. Thus, a non-negative measurable function f is integrable on E if and only if

$$\lim_{n\to\infty} \int_{E_n} [f]_n(x)\, dx < \infty$$

for some (equivalently, for all) finite monotonic cover $\{E_n\}$ of $m(E)$.

(ii) In particular, if $m(E) < \infty$ and the non-negative measurable function f is bounded on E, then $\int_E f(x)\, dx$ defined above coincides with that given in the last section.

Indeed, we can choose $E_n = E$ for all $n \in \mathbb{N}$ and $N = \sup_{x \in E} f(x)$.

Example. Consider the function f on $(0, \infty)$ defined by

$$f(x) := \begin{cases} x^{-2}, & x \geq 1; \\ x^{-1/2}, & 0 < x < 1. \end{cases}$$

We have

$$(L) \int_0^\infty f(x)\, dx = 3.$$

Proof. Let $E_n = [\frac{1}{n^2}, n]$ and $M_n = n$ for $n \in \mathbb{N}$. Then

$$\begin{aligned} \int_{E_n} [f]_{M_n}(x)\, dx &= \int_{n^{-2}}^1 [f]_{M_n}(x)\, dx + \int_1^n [f]_{M_n}(x)\, dx \\ &= \int_{n^{-2}}^1 \frac{1}{\sqrt{x}}\, dx + \int_1^n \frac{1}{x^2}\, dx \\ &= 3 - 3/n \to 3, \quad n \to \infty. \end{aligned}$$

$\square$

Integrals for general real-valued functions.

Recall that each real-valued function defined on $E \subset \mathbb{R}$ has a decomposition

$$f = f_+ - f_-,$$

where

$$f_+(x) := \begin{cases} f(x), & f(x) \geq 0; \\ 0, & f(x) < 0, \end{cases}$$

and

$$f_-(x) := \begin{cases} -f(x), & f(x) \leq 0; \\ 0, & f(x) > 0. \end{cases}$$

Clearly, both f_+ and f_- are non-negative functions and $|f| = f_+ + f_-$.

Definition. Let $E \subset \mathbb{R}$ be Lebesgue measurable. A Lebesgue measurable function f on E is said to be **Lebesgue integrable** on E if f_+ and f_- are both Lebesgue integrable on E. In this case, the **Lebesgue integral** of f is defined by

$$\int_E f(x)\,dx := \int_E f_+(x)\,dx - \int_E f_-(x)\,dx.$$

Lemma. Let f be non-negative and Lebesgue integrable on E and g Lebesgue measurable on E. If $|g| \leq f$, then g is Lebesgue integrable on E.

Proof. It suffices to show that g_+ and g_- are both integrable. Let $\{E_n\}$ be a finite monotonic cover of the measure of E. For each $n \in \mathbb{N}$, it follows that

$$\int_{E_n} [g_\pm]_n(x)\,dx \leq \int_{E_n} [f]_n(x)\,dx \leq \int_E f(x)\,dx.$$

Thus,

$$\lim_{n\to\infty}\int_{E_n}[g_\pm]_n(x)\,dx \le \int_E f(x)\,dx < \infty,$$

which implies that g_+ and g_- are both integrable. $\square$

Remark. This lemma reads that if a measurable function g can be dominated by an integrable function, then g is integrable.

Theorem. (**Finite additivity**) Let E, E_1 and E_2 are all Lebesgue measurable subsets of $\mathbb{R}$, $E = E_1 \cup E_2, E_1 \cap E_2 = \emptyset$, and let f be Lebesgue measurable on E. Then f is Lebesgue integrable on E if and only if f is Lebesgue integrable on both E_1 and E_2. Moreover,

$$\int_E f(x)\,dx = \int_{E_1} f(x)\,dx + \int_{E_2} f(x)\,dx.$$

Proof. (I) Suppose that $f \ge 0$. Let $\{E_n\}$ be a finite monotonic cover of the measure of E. Clearly, $\{E_n \cap E_i\}$ is a finite monotonic cover of the measure of E_i $(i = 1, 2)$. For each $n \in \mathbb{N}$, we have

$$\begin{aligned}
\int_{E_n}[f]_n(x)\,dx &= \int_{E_n\cap E_1}[f]_n(x)\,dx + \int_{E_n\cap E_2}[f]_n(x)\,dx \\
&= \int_{E_n\cap E_1}[f]_n|_{E_1}(x)\,dx + \int_{E_n\cap E_2}[f]_n|_{E_2}(x)\,dx \\
&= \int_{E_n\cap E_1}[f|_{E_1}]_n(x)\,dx + \int_{E_n\cap E_2}[f|_{E_2}]_n(x)\,dx \\
&\le \int_{E_1} f|_{E_1}(x)\,dx + \int_{E_2} f|_{E_2}(x)\,dx \\
&= \int_{E_1} f(x)\,dx + \int_{E_2} f(x)\,dx.
\end{aligned}$$

Thus, if f is integrable on E_1 and E_2, then f is integrable on E.

Conversely, let f be integrable on E. Then

$$\begin{aligned}\int_{E_n\cap E_i}[f|_{E_i}]_n(x)\,dx &\leq \int_{E_i}[f|_{E_i}]_n(x)\,dx\\ &\leq \int_E f(x)\,dx \quad (n\in\mathbb{N})\end{aligned}$$

for $i=1,2$. Thus, $f|_{E_1}$ and $f|_{E_2}$ are integrable on E_1 and E_2, respectively. That is, f is integrable on E_1 and E_2.

In addition, taking $n\to\infty$ in both sides of

$$\int_{E_n}[f]_n(x)\,dx = \int_{E_n\cap E_1}[f]_n(x)\,dx + \int_{E_n\cap E_2}[f]_n(x)\,dx$$

yields the equality desired.

(II) For general function f (not necessarily non-negative), consider $f=f_+-f_-$. It follows from the step I that

$$\begin{aligned}\int_E f(x)\,dx &= \int_E f_+(x)\,dx - \int_E f_-(x)\,dx\\ &= \int_{E_1} f_+(x)\,dx + \int_{E_2} f_+(x)\,dx\\ &\quad - \int_{E_1} f_-(x)\,dx - \int_{E_2} f_-(x)\,dx\\ &= \int_{E_1} f(x)\,dx + \int_{E_2} f(x)\,dx.\end{aligned}$$

□

Theorem. (**Linear property**) Let f and g be two Lebesgue integrable functions on E. Then for all $a, b \in \mathbb{R}$, $af + bg$ is Lebesgue integrable on E and

$$\int_E (af(x) + bg(x))\,dx = a\int_E f(x)\,dx + b\int_E g(x)\,dx.$$

Proof. It suffices to prove

$$\int_E af(x)\,dx = a\int_E f(x)\,dx \tag{0.8}$$

and

$$\int_E (f(x) + g(x))\,dx = \int_E f(x)\,dx + \int_E g(x)\,dx. \tag{0.9}$$

The following method is standard one followed by the definition of integrals.

Step I. Non-negative case. Suppose that $f \geq 0$. We first show (0.8) holds. The case that $a = 0$ is trivial. For $a > 0$, notice that $[af]_N = a[f]_{N/a}$, we then have

$$\begin{aligned}\int_E af(x)\,dx &= \lim_{n\to\infty}\int_{E_n}[af]_N(x)\,dx\\ &= \lim_{n\to\infty} a\int_{E_n}[f]_{N/a}(x)\,dx = a\int_E f(x)\,dx.\end{aligned}$$

For $a < 0$, notice that

$$\begin{aligned}&\int_E (-f)(x)\,dx = \int_E (-f)_+(x)\,dx - \int_E (-f)_-(x)\,dx\\ &= \int_E f_-(x)\,dx - \int_E f_+(x)\,dx = -\int_E f(x)\,dx,\end{aligned}$$

we then have

$$\begin{aligned}\int_E af(x)\,dx &= \int_E -(-a)f(x)\,dx\\ &= -\int_E (-a)f(x)\,dx = a\int_E f(x)\,dx.\end{aligned}$$

This proves (0.8). On the other hand, notice that

$$[f+g]_N \le [f]_N + [g]_N \le [f+g]_{2N} \tag{0.10}$$

implies that $f+g$ is integrable if and only if f and g are both integrable. Moreover, (0.9) follows from (0.10), immediately.

Step II. General case. Notice that, in the step I, we have proved in fact that

$$\begin{aligned}&\int_E (a_1 f_1 + a_2 f_2 + \cdots + a_n f_n)(x)\,dx\\ =&a_1\int_E f_1(x)\,dx + a_2\int_E f_2(x)\,dx + \cdots + a_n\int_E f_n(x)\,dx\end{aligned}$$

for all $f_1, \cdots, f_n \ge 0$ and $a_1, \cdots, a_n \in \mathbb{R}$, where $n \in \mathbb{N}$ arbitrary. Thus,

$$\begin{aligned}&\int_E (af+bg)(x)\,dx\\ =&\int_E (af_+ - af_- + bg_+ - bg_-)(x)\,dx\\ =&a\int_E f_+(x)\,dx - a\int_E f_-(x)\,dx\\ &+b\int_E g_+(x)\,dx - b\int_E g_-(x)\,dx\\ =&a\int_E f(x)\,dx + b\int_E g(x)\,dx.\end{aligned}$$

$\square$

Theorem. Let f and g be two measurable functions on E. The following statements hold.

(i) If f is integrable on E, then $\int_E f(x)\,dx = 0$ whenever $m(E) = 0$.

(ii) If f is integrable on E and $E_1 \subset E$ with $m(E_1) = 0$, then f is integrable on E_1 and

$$\int_{E\setminus E_1} f(x)\,dx = \int_E f(x)\,dx.$$

(iii) If $f = 0$ (a.e.) on E, then f is integrable and

$$\int_E f(x)\,dx = 0.$$

(iv) If f is integrable on E and $f \geq 0$ (a.e.) on E, then

$$\int_E f(x)\,dx \geq 0.$$

(v) If $f \geq 0$ (a.e.) on E and $\int_E f(x)\,dx = 0$, then $f = 0$ (a.e.) on E.

(vi) (**Monotonicity**) If f and g are both integrable and $f \leq g$ (a.e.), then

$$\int_E f(x)\,dx \leq \int_E g(x)\,dx.$$

(vii) (**Absolute integrability**) If f is integrable on E, then $|f|$ is also integrable on E, moreover,

$$\left| \int_E f(x)\,dx \right| \leq \int_E |f|(x)\,dx.$$

(viii) If f is integrable on E, then f is integrable on each measurable subset $F \subset E$. Moreover,

$$\left| \int_F f(x)\,dx \right| \leq \int_F |f|(x)\,dx \leq \int_E |f|(x)\,dx.$$

Proof. We only give a proof of (v), the others are standard and very easy, so omitted.

(v) Let $f \geq 0$ (a.e.) on E and $\int_E f(x)\,dx = 0$. Then

$$\begin{aligned} 0 &= \int_E f(x)\,dx = \lim_{n\to\infty} \int_{E_n} [f]_n(x)\,dx \geq \int_{E_n} [f]_n(x)\,dx \\ &= \int_{E_n(f<\frac{1}{n})} [f]_n(x)\,dx + \int_{E_n(f\geq\frac{1}{n})} [f]_n(x)\,dx \\ &\geq \int_{E_n(f\geq\frac{1}{n})} [f]_n(x)\,dx \geq \frac{1}{n} m\Big(E_n\big(f \geq \frac{1}{n}\big)\Big), \end{aligned}$$

so that $m(E_n(f \geq \frac{1}{n})) = 0$ for all $n \in \mathbb{N}$.

On the other hand, we show

$$E(f > 0) = \bigcup_{n=1}^{\infty} E_n\Big(f \geq \frac{1}{n}\Big).$$

Clearly, it suffices to prove

$$E(f > 0) \subset \bigcup_{n=1}^{\infty} E_n\Big(f \geq \frac{1}{n}\Big).$$

Indeed, let $x \in E(f > 0)$. Then $f(x) = c > 0$. Take $n \in \mathbb{N}$ such that $c \geq \frac{1}{n}$. Then

$$x \in E(f \geq \frac{1}{n}) = \bigcup_{m=1}^{\infty} E_m\Big(f \geq \frac{1}{n}\Big),$$

so that there is $m \in \mathbb{N}$ such that $x \in E_m(f \geq \frac{1}{n})$. Denote $N := \max\{m, n\}$. Then

$$x \in E_m\Big(f \geq \frac{1}{n}\Big) \subset E_m\Big(f \geq \frac{1}{N}\Big) \subset E_N\Big(f \geq \frac{1}{N}\Big).$$

This proves the inclusion relation desired.

Therefore,

$$\begin{aligned} m(E(f>0)) =& m\Big(\bigcup_{n=1}^{\infty} E_n(f \geq \frac{1}{n})\Big) \\ & \leq \sum_{n=1}^{\infty} m(E_n(f \geq \frac{1}{n})) = 0. \end{aligned}$$

□

Remark. From (vi) above, we observe that f **is integrable on** E **if and only if** $|f|$ **is integrable on** E, however, it is not true for generalized Riemann integrals. For example, consider the function f on $(0, \infty)$ defined by

$$f(x) := \frac{\sin x}{x}, \quad 0 < x < \infty.$$

Then f is generalized Riemann integrable, however, $|f|$ is not generalized Riemann integrable. In addition, f is not Lebesgue integrable. Thus, **the new integral is an extension of the proper Riemann integral (or, the improper Riemann integral of non-negative functions), not the general improper Riemann integral**.

Theorem. (**Complete continuity**) Let f be integrable on E. Then for each $\epsilon > 0$, there is a $\delta > 0$ such that

$$\left| \int_{E'} f(x)\, dx \right| < \epsilon$$

for every measurable subset $E' \subset E$ with $m(E') < \delta$.

Proof. Notice that f integrable if and only if $|f|$ is integrable. By the definition of integrals,

$$\lim_{n\to\infty} \int_{E_n} [|f|]_n(x)\, dx = \int_E |f(x)|\, dx < \infty,$$

where $\{E_n\}$ is a finite monotonic cover of $m(E)$. Let $\epsilon > 0$ arbitrary. Then there is $N \in \mathbb{N}$ such that

$$\int_E |f(x)|\, dx - \int_{E_N} [|f|]_N(x)\, dx < \frac{\epsilon}{2}.$$

Take $\delta = \frac{\epsilon}{2(N+1)}$. Then for each $E' \subset E$ with $m(E') < \delta$ we have

$$\begin{aligned}
&\left| \int_{E'} f(x)\, dx \right| \leq \int_{E'} |f|(x)\, dx \\
=& \int_{E'} (|f|(x) - [|f|]_N(x))\, dx + \int_{E'} [|f|]_N(x)\, dx \\
\leq& \int_{E} (|f|(x) - [|f|]_N(x))\, dx + N \cdot m(E') \\
\leq& \int_{E} |f(x)|\, dx - \int_{E_N} [|f|]_N(x)\, dx + \frac{N}{2(N+1)}\epsilon \\
<& \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

□

Remark. $\epsilon \leftrightarrow$ approximation $\leftrightarrow$ lim.

Let f be integrable on E and let E_1, E_2 be two measurable subsets of E with $E = E_1 \cup E_2$ and $E_1 \cap E_2 = \emptyset$. By the monotonicity, f is integrable on E_1 and E_2. In addition, we have

$$\int_E f(x)\,dx = \int_{E_1} f(x)\,dx + \int_{E_2} f(x)\,dx.$$

This is the **finite additivity** so called. Further on, we have the following countable additivity.

Theorem. (**Countable additivity**) Let f be a measurable function on E, and let $\{E_n\}$ be a sequence of measurable subsets of E satisfying $E_i \cap E_j = \emptyset$ for $i \neq j$ and $E = \bigcup_{n=1}^{\infty} E_n$. Then f is integrable on E if and only if

(i) f is integrable on E_n for all $n \in \mathbb{N}$;

(ii) $\sum_{n=1}^{\infty} \int_{E_n} |f(x)|\,dx < \infty$.

Moreover, if f is integrable on E then

$$\int_E f(x)\,dx = \sum_{n=1}^{\infty} \int_{E_n} f(x)\,dx. \tag{0.11}$$

Proof. Necessity. Let f be integrable on E. Notice that f is integrable if and only if $|f|$ is integrable, so that

$$\int_{E_n} |f|(x)\,dx \leq \int_E |f|(x)\,dx < \infty$$

for all $n \in \mathbb{N}$. This yields (i). On the other hand, by the finite additivity of integrals, we have

$$\begin{aligned}\sum_{n=1}^{m} \int_{E_n} |f|(x)\,dx &= \int_{\cup_{n=1}^{m} E_n} |f|(x)\,dx \\ &\leq \int_E |f|(x)\,dx < \infty\end{aligned}$$

for all $m \in \mathbb{N}$. By letting $m \to \infty$ in both sides of the inequality above we obtain (ii). In fact, we have

$$\sum_{n=1}^{\infty} \int_{E_n} |f|(x)\,dx \le \int_E |f|(x)\,dx. \tag{0.12}$$

Sufficiency. Let $\{F_n\}$ be a finite monotonic cover of $m(E)$. We will prove that $|f|$ is integrable. By definition, it suffices to show the limit

$$\lim_{n\to\infty} \int_{F_n} [|f|]_n(x)\,dx < \infty.$$

Notice that $\{F_n'\}$ with

$$F_n' := F_n \cap \left(\bigcup_{k=1}^{n} E_k \right) = \bigcup_{k=1}^{n} (F_n \cap E_k) \quad (n \in \mathbb{N})$$

is a finite monotonic cover of $m(E)$ as well, then by the finite additivity of integrals of bounded functions on sets with finite measures, we have

$$\begin{aligned} \int_{F_n'} [|f|]_n(x)\,dx &= \sum_{k=1}^{n} \int_{F_n \cap E_k} [|f|]_n(x)\,dx \\ &\le \sum_{k=1}^{n} \int_{E_k} |f|(x)\,dx \le \sum_{k=1}^{\infty} \int_{E_k} |f|(x)\,dx. \end{aligned}$$

By letting $n \to \infty$ we obtain the inequality desired. In fact, we have

$$\int_E |f|(x)\,dx \le \sum_{n=1}^{\infty} \int_{E_n} |f|(x)\,dx. \tag{0.13}$$

In addition, it follows from (0.12) and (0.13) that

$$\sum_{n=1}^{\infty} \int_{E_n} |f|(x)\,dx = \int_E |f|(x)\,dx. \tag{0.14}$$

Finally, we prove the equality (0.11). Indeed, it follows from the finite additivity property and (0.14) that

$$
\begin{aligned}
&\left| \int_E f(x)\,dx - \sum_{n=1}^{m} \int_{E_n} f(x)\,dx \right| \\
=&\left| \int_E f(x)\,dx - \int_{\bigcup_{n=1}^m E_n} f(x)\,dx \right| = \left| \int_{E\setminus\bigcup_{n=1}^m E_n} f(x)\,dx \right| \\
\leq& \int_{E\setminus\bigcup_{n=1}^m E_n} |f|(x)\,dx = \int_E |f|(x)\,dx - \int_{\bigcup_{n=1}^m E_n} |f|(x)\,dx \\
=& \sum_{n=m+1}^{\infty} \int_{E_n} |f|(x)\,dx \to 0, \quad m \to \infty.
\end{aligned}
$$

□

Theorem. (**Relation of integrable function and continuous function**) Let f be integrable on $E = [a, b]$. Then for each $\epsilon > 0$, there is a function f_ϵ continuous on $[a, b]$ such that

$$(L)\int_a^b |f(x) - f_\epsilon(x)|\, dx < \epsilon.$$

Proof. Let $\epsilon > 0$ and let $\{E_n\}$ be a finite monotonic cover of $m(E)$. By definition, we have

$$\begin{aligned}\int_E \big|f(x) - [f]_n(x)\big|\, dx &= \int_E (f(x) - [f]_n(x))\, dx \\ &= \int_E f(x)\, dx - \int_E [f]_n(x)\, dx \\ &\leq \int_E f(x)\, dx - \int_{E_n} [f]_n(x)\, dx \to 0\end{aligned}$$

as $n \to \infty$, so that there is $N \in \mathbb{N}$ such that

$$\int_E \big|f(x) - [f]_N(x)\big|\, dx < \frac{\epsilon}{2}.$$

Take $\delta = \frac{\epsilon}{4N}$. For the function $[f]_N$, by the Lusin's theorem, there are measurable subset $E_\delta \subset E$ with $m(E \setminus E_\delta) < \delta$ and continuous function g on E with $|g| \leq N$ such that $[f]_N = g$ on E_δ. g is just the function f_ϵ we desired. Indeed,

$$\begin{aligned}\int_E |f(x) - g(x)|\, dx \leq &\int_E \big|f(x) - [f]_N(x)\big|\, dx \\ &+ \int_E \big|[f]_N(x) - g(x)\big|\, dx \\ <&\frac{\epsilon}{2} + \int_{E \setminus E_\delta} \big|[f]_N(x) - g(x)\big|\, dx \\ \leq&\frac{\epsilon}{2} + 2N \cdot m(E \setminus E_\delta) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

□

§3. Limit theorems

In this section, we introduce three theorems of extreme importance in the theory of analysis: **Lebesgue dominated convergence theorem**, **Levi lemma** and **Fatou lemma**.

Let f and g be integrable functions on $E \subset \mathbb{R}$. Recall that the linear property of the integral reads that

$$\int_{\mathbb{R}} cf(x)\,dx = c\int_{\mathbb{R}} f(x)\,dx$$

for $c \in \mathbb{R}$ and

$$\int_{\mathbb{R}} (f(x) + g(x))\,dx = \int_{\mathbb{R}} f(x)\,dx + \int_{\mathbb{R}} g(x)\,dx.$$

The first equality shows that the order of operations of scalar-multiplication and integral can be exchanged and the second equality implies that the operations of addition and integral can be exchanged, unconditionally.

However, it is not true for operations of the limit and integral. See the following

Counterexample. Let $f_n = \chi_{[n,n+1)}$ for $n \in \mathbb{N}$. It is clear that

$$\int_{\mathbb{R}} f_n(x)\,dx = 1, \quad n = 1, 2, \cdots .$$

Moreover, for each $x \in \mathbb{R}$ fixed, $\lim_{n\to\infty} f_n(x) = 0$. Thus,

$$1 = \lim_{n\to\infty} \int_{\mathbb{R}} f_n(x)\,dx \neq \int_{\mathbb{R}} \lim_{n\to\infty} f_n(x)\,dx = 0.$$

Lebesgue dominated convergence theorem

Question. Under what conditions, the order of operations lim and $\int$ can be exchanged?

To solve this problem, we need the following "dominant function" so called.

Dominant function. Let $E \subset \mathbb{R}$ be a measurable set and $\{f_n\}$ a sequence of measurable functions on E, and let F be a non-negative measurable function on E. We say that $\{f_n\}$ **is dominated by** F (or F is a **dominant function** of $\{f_n\}$) if $|f_n| \leq F$ (a.e.) for all $n \in \mathbb{N}$.

Theorem. (**Lebesgue dominated convergence**) Let $\{f_n\}$ be a sequence of measurable functions on E dominated by an **integrable** function F. If f_n converges almost everywhere to a measurable function f on E, then f is integrable on E and

$$\lim_{n\to\infty} \int_E f_n(x)\,dx = \int_E f(x)\,dx.$$

Remark. The Lebesgue dominated convergence theorem reads that

$$\lim_{n\to\infty} \int_E = \int_E \lim_{n\to\infty}$$

whenever $\{f_n\}$ can be **dominated by an integrable function**.

Proof. Notice that $|f| \leq F$ (a.e.), so f is integrable on E due to the integrability of F. It suffices to show

$$\lim_{n\to\infty} \int_E (f_n(x) - f(x))\,dx = 0. \tag{0.15}$$

To show (0.15) hold, we need an estimate for the integral

$$\int_E |f_n(x) - f(x)|\,dx.$$

Step I. Let $m(E) < \infty$ and let $\epsilon > 0$ arbitrary. Since F is integrable on E, by the complete continuity of integral, there is a $\delta > 0$ such that

$$\int_{E'} F(x)\,dx < \frac{\epsilon}{4} \tag{0.16}$$

for each measurable subset $E' \subset E$ with $m(E') < \delta$. On the other hand, since $f_n \to f$ (a.e.), by Egoroff's theorem, for the δ above, there is a measurable subset $E_\delta \subset E$ with $m(E \setminus E_\delta) < \delta$ such that $f_n \rightrightarrows f$ on E_δ. Thus, for the ϵ above, there is $N \in \mathbb{N}$ such that

$$\sup_{x \in E_\delta} |f_n(x) - f(x)| \leq \frac{\epsilon}{2m(E)+1} \tag{0.17}$$

for all $n \geq N$. Therefore, it follows from (0.16) and (0.17) that

$$\begin{aligned}
&\left| \int_E (f_n(x) - f(x))\,dx \right| \leq \int_E |f_n(x) - f(x)|\,dx \\
&= \int_{E \setminus E_\delta} |f_n(x) - f(x)|\,dx + \int_{E_\delta} |f_n(x) - f(x)|\,dx \\
&\leq \int_{E \setminus E_\delta} 2F(x)\,dx + \int_{E_\delta} \frac{\epsilon}{2m(E)+1}\,dx \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

for all $n \geq N$. This proves (0.15).

Step II. For general E (not necessarily $m(E) < \infty$), it is a little complicated. Let $\epsilon > 0$ arbitrary, and let $\{E_k\}$ be a finite monotonic cover of $m(E)$. By the definition of integral,

$$\lim_{k \to \infty} \int_{E_k} [F]_k(x)\,dx = \int_E F(x)\,dx,$$

so that there is $K \in \mathbb{N}$ such that

$$\int_E F(x)\,dx - \int_{E_k} [F]_k(x)\,dx < \frac{\epsilon}{4}$$

for all $k \geq K$. Thus,

$$
\begin{aligned}
&\sup_{n\in\mathbb{N}} \int_{E\setminus E_k} |f_n(x) - f(x)|\, dx \leq 2 \int_{E\setminus E_k} F(x)\, dx \\
&=2\left(\int_E F(x)\, dx - \int_{E_k} F(x)\, dx \right) \\
&\leq 2\left(\int_E F(x)\, dx - \int_{E_k} [F]_k(x)\, dx \right) \\
&< \frac{\epsilon}{2}
\end{aligned} \tag{0.18}
$$

for all $k \geq K$. Consider the **special value** $k = K$. Notice that $m(E_K) < \infty$. By the step I, for the ϵ above, there is $N \in \mathbb{N}$ such that

$$
\int_{E_K} |f_n(x) - f(x)|\, dx < \frac{\epsilon}{2} \tag{0.19}
$$

for all $n \geq N$. Then it follows from (0.18) and (0.19) that

$$
\begin{aligned}
\int_E |f_n(x) - f(x)|\, dx = &\int_{E\setminus E_K} |f_n(x) - f(x)|\, dx \\
&+ \int_{E_K} |f_n(x) - f(x)|\, dx < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}
$$

for all $n \geq N$. This completes the proof. $\square$

Levi lemma. (**Monotonic convergence**) Let $\{f_n\}$ be an increasing sequence of integrable functions on E. If

$$\sup_{n\in\mathbb{N}} \int_E f_n(x)\, dx := M < \infty,$$

then $f_n \to f$ (a.e.) for some function f integrable on E. Moreover,

$$\lim_{n\to\infty} \int_E f_n(x)\, dx = \int_E f(x)\, dx.$$

Proof. Step I. Suppose that $f_n \geq 0$ for all $n \in N$. Notice that

$$f_1 \leq f_2 \leq \cdots \leq f_n \leq \cdots,$$

by denoting $f^*(x) := \lim_{n\to\infty} f_n(x)$ for $x \in E$, then f^* is a function from E to $[0, \infty]$.

We first show that $m(E(f^* = \infty)) = 0$. Let $\{E_n\}$ be a finite monotonic cover of $m(E)$. Notice that

$$E(f^* = \infty) = \bigcup_{n=1}^{\infty} E_n(f^* = \infty) = \lim_{n\to\infty} E_n(f^* = \infty),$$

since $\{E_n(f^* = \infty)\}$ is an increasing sequence. It suffices to show $m(E_n(f^* = \infty)) = 0$ for all $n \in \mathbb{N}$. Indeed, for $N > 0$ arbitrary, notice that

$$0 \leq [f_1]_N \leq [f_2]_N \leq \cdots \leq [f_n]_N \leq \cdots \to [f^*]_N,$$

then by the Lebesgue's dominated convergence theorem, we obtain

$$\begin{aligned}\int_{E_N} [f^*]_N(x)\, dx &= \lim_{n\to\infty} \int_{E_N} [f_n]_N(x)\, dx \\ &\leq \lim_{n\to\infty} \int_E f_n(x)\, dx \leq M.\end{aligned} \tag{0.20}$$

For $f^*(x) > N$, $[f^*]_N = N$, so that $E([f^*]_N = N)$ is measurable, and from (0.20) it follows that

$$\begin{aligned} m(E_n(f^* \geq N)) =& m(E_n([f^*]_N = N)) \\ =& \frac{1}{N} \int_{E_n([f^*]_N = N)} [f^*]_N(x)\, dx \\ \leq& \frac{1}{N} \int_{E_N} [f^*]_N(x)\, dx \leq \frac{M}{N}, \quad N \geq n, \end{aligned}$$

which implies that $\lim_{N\to\infty} m(E_n(f^* \geq N)) = 0$. Thus,

$$\begin{aligned} m(E_n(f^* = \infty)) =& m\Bigg(\bigcap_{N=1}^{\infty} E_n(f^* \geq N) \Bigg) \\ =& m\Bigg(\lim_{N\to\infty} E_n(f^* \geq N) \Bigg) \\ =& \lim_{N\to\infty} m\Bigg(E_n(f^* \geq N) \Bigg) = 0. \end{aligned}$$

Next, define a function f by

$$f(x) := \begin{cases} f^*(x), & f^*(x) < \infty; \\ 0, & f^*(x) = \infty. \end{cases}$$

Then $f = f^*$ (a.e.), so that $f_n \to f$ (a.e.) as $n \to \infty$. Further on, it follows from (0.20) that

$$\int_{E_N} [f]_N(x)\, dx = \int_{E_N} [f^*]_N(x)\, dx \leq \int_E [f^*]_N(x)\, dx \leq M,$$

while $\int_{E_N} [f]_N(x)\, dx$ is increasing with respect to the index $N \in \mathbb{N}$, so that the limit

$$\lim_{N\to\infty} \int_{E_N} [f]_N(x)\, dx \leq M < \infty.$$

This proves f is integrable on E.

Finally, notice that f is a dominant function of f_n, again by the Lebesgue's dominated convergence theorem, we obtain

$$\lim_{n\to\infty}\int_E f_n(x)\,dx = \int_E f(x)\,dx.$$

Step II. For general sequence $\{f_n\}$, consider

$$g_n := f_n - f_1, \quad n \in \mathbb{N}.$$

It is clear that g_n is non-negative for all $n \in \mathbb{N}$. Then by the step I, $\{g_n\}$ converges almost everywhere to an integrable g on E, or equivalently, $\{f_n\}$ converges almost everywhere to an integrable f on E, and

$$\lim_{n\to\infty}\int_E g_n(x)\,dx = \lim_{n\to\infty}\left(\int_E f_n(x)\,dx - \int_E f_1(x)\,dx\right)$$
$$= \int_E \lim_{n\to\infty} g_n(x)\,dx = \int_E \left(\lim_{n\to\infty} f_n(x) - f_1(x)\right) dx.$$

Therefore, the limit $\lim_{n\to\infty}\int_E f_n(x)\,dx$ exists and

$$\lim_{n\to\infty}\int_E f_n(x)\,dx = \int_E f(x)\,dx.$$

This completes the proof. □

By considering the part sum $\sum_{k=1}^n u_k := f_n$, we obtain the following series version of Levi's lemma, immediately.

Levi lemma II. Let $\{u_n\}$ be a sequence of non-negative integrable functions on E. If $\sum_{n=1}^\infty \int_E u_n(x)\,dx < \infty$, then $\sum_{n=1}^\infty u_n = u$ (a.e.) for some integrable function u on E, and

$$\int_E u(x)\,dx = \sum_{n=1}^\infty \int_E u_n(x)\,dx.$$

Counterexample. Consider the sequence of functions f_n defined on $[0,1]$ by

$$f_n(x) := \begin{cases} \frac{1}{x}\left|\sin\frac{1}{x}\right|, & \frac{1}{n} \le x \le n; \\ 0, & 0 \le x < \frac{1}{n}. \end{cases}$$

It is clear that $\{f_n\}$ is increasing and

$$f_n(x) \to f(x) = \begin{cases} \frac{1}{x}\left|\sin\frac{1}{x}\right|, & 0 < x \le 1; \\ 0, & x = 0. \end{cases}$$

However, f is not integrable on $[0,1]$. Observe that

$$\lim_{n\to\infty} (L)\int_0^1 f_n(x)\,dx = \infty$$

in this example. Thus, the condition

$$\text{“}\sup_{n\in\mathbb{N}} \int_E f_n(x)\,dx < \infty\text{”}$$

in Levi's lemma is **indispensable**.

Counterexample II. For $n \in \mathbb{N}$, define $u_n : \mathbb{R} \to \mathbb{R}$ by

$$u_n := \chi_{[n,n+1)} - \chi_{[n+1,n+2)}.$$

It is clear that

$$\sum_{n=1}^{\infty} \int_{\mathbb{R}} u_n(x)\,dx = 0 < \infty$$

and $\sum_{k=1}^{n} u_k(x)$ converges in point-wise sense to an integrable function u:

$$u(x) = \begin{cases} 1, & x \in [1,2); \\ 0, & \text{otherwise}. \end{cases}$$

However,

$$\int_{\mathbb{R}} \sum_{n=1}^{\infty} u_n(x)\,dx = 1 \ne 0 = \sum_{n=1}^{\infty} \int_{\mathbb{R}} u_n(x)\,dx.$$

Fatou lemma.

Recall that the sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$ defined by $f := \chi_{[n,n+1)}$ can not be dominated by an integrable on $\mathbb{R}$. Further on, we have

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n(x)\,dx \neq \int_{\mathbb{R}} \lim_{n\to\infty} f_n(x)\,dx.$$

In fact,

$$\int_{\mathbb{R}} \lim_{n\to\infty} f_n(x)\,dx = 0 < 1 = \lim_{n\to\infty} \int_{\mathbb{R}} f_n(x)\,dx.$$

This example shows that it is possible for some **non-negative** sequence $\{f_n\}$ that

$$\int_{\mathbb{R}} \lim_{n\to\infty} f_n(x)\,dx < \lim_{n\to\infty} \int_{\mathbb{R}} f_n(x)\,dx.$$

However, **the inverse inequality can not be happen**, which is stated in detail in the following

Fatou lemma. Let $\{f_n\}$ be a sequence of integrable functions on E. If $f_n \geq h$ (a.e.) for some function h integrable on E and

$$\varliminf_{n\to\infty} \int_E f_n(x)\,dx < \infty,$$

then the function $\varliminf_{n\to\infty} f_n$ is integrable on E and

$$\int_E \varliminf_{n\to\infty} f_n(x)\,dx \leq \varliminf_{n\to\infty} \int_E f_n(x)\,dx.$$

Proof. For $x \in E$, we write

$$F_{n,m}(x) := \min\{f_n(x), f_{n+1}(x), \cdots, f_{n+m}(x)\}, \quad n, m \in \mathbb{N}.$$

Then

$$\lim_{n\to\infty} \lim_{m\to\infty} F_{n,m}(x) := \lim_{n\to\infty} \inf_{m\geq n} f_m(x) = \varliminf_{n\to\infty} f_n(x).$$

Fix $n \in \mathbb{N}$. Notice that $F_{n,m} \geq h$ for all $m \in \mathbb{N}$ and $\{F_{n,m}\}_{m=1}^{\infty}$ is decreasing, so that by Levi's lemma $F_{n,m} \to F_n$ (a.e.) as $m \to \infty$ for some integrable function F_n and

$$\lim_{m\to\infty} \int_E F_{n,m}(x)\,dx = \int_E \lim_{m\to\infty} F_{n,m}(x)\,dx = \int_E F_n(x)\,dx.$$

On the other hand, by letting $m \to \infty$ in the following inequality

$$\int_E F_{n,m}(x)\,dx \leq \min\left\{ \int_E f_n(x)\,dx, \cdots, \int_E f_{n+m}(x)\,dx \right\},$$

we obtain

$$\int_E F_n(x)\,dx \leq \inf_{m\geq n} \int_E f_m(x)\,dx,$$

so that

$$\begin{aligned}\sup_{n\geq 1} \int_E F_n(x)\,dx &\leq \sup_{n\geq 1} \inf_{m\geq n} \int_E f_m(x)\,dx \\ &= \varliminf_{n\to\infty} \int_E f_n(x)\,dx < \infty.\end{aligned}$$

Notice that $\{F_n\}$ is increasing, again by the Levi lemma, $F_n \to F$ (a.e.) for some integrable function F and

$$\begin{aligned}\int_E \varliminf_{n\to\infty} f_n(x)\,dx &= \int_E \lim_{n\to\infty} F_n(x)\,dx \\ &= \lim_{n\to\infty} \int_E F_n(x)\,dx \\ &= \sup_{n\geq 1} \int_E F_n(x)\,dx \leq \varliminf_{n\to\infty} \int_E f_n(x)\,dx.\end{aligned}$$

This completes the proof. □

Remark. Let $\{f_n\}$ be a sequence of non-negative integrable functions on E and $f_n \to f$ (a.e.) for some integrable function f on E. If the limit

$$\lim_{n\to\infty}\int_E f_n(x)\,dx < \infty,$$

then by the Fatou lemma we observe that

$$\int_E \lim_{n\to\infty} f_n(x)\,dx \le \lim_{n\to\infty}\int_E f_n(x)\,dx, \tag{0.21}$$

as showed in the beginning of this section. Notice that **the non-negativity of $\{f_n\}$ is indispensable** in (0.21), see the following

Counterexample. Consider functions

$$f_n := -\chi_{[n,n+1)}, \quad n \in \mathbb{N}.$$

It is clear that

$$\int_{\mathbb{R}} \lim_{n\to\infty} f_n(x)\,dx = 0 > -1 = \lim_{n\to\infty}\int_{\mathbb{R}} f_n(x)\,dx.$$