FUNDAMENTAL ALGEBRA & ANALYSIS

Compiled by **Huayi Chen**[I],

&

Edited by **Jiete Xue**[II],

– –

TAs: **Jiedong Jiang, Yijun Yuan, Chunhui Liu**[III, IV, V], and

◇

Substitute Professor: **Yigeng Zhao, Weronica Aniela Czerniawska**[VI, VII]

[I]*Department of Mathematics (ITS), School of Science, Westlake University*
[II]*Undegraduate β Collage, Westlake University*
[I]`chenhuayi@westlake.edu.cn`
[II]`xuejiete@westlake.edu.cn`

Contents

Chapter 1

Basic Logic

1.1 Statement

Definition 1.1.1 We call statement a declarative sentence that is either true or false, but not both (it can be potential).

Example 1.1.2 "$2 > 1$" (True) "$1 < 0$" (False)
If we specify the value of x , then "$x > 2$" becomes a statement, otherwise it is not a statement.

Definition 1.1.3 In a mathematical theory,
axiom refer to statements that accepted to be true without justification.
theorem refer to statements that are proved by assuming axioms.
proposition refer to the statements that are either easy or not used many times.
corollary refer to direct consequence of a theorem.

1.2 Negation

Definition 1.2.1 Let p be a statement, then the negation of p is denoted by $\neg p$, which is a statement that is true if and only if p is false. In other words, p and $\neg p$ has opposite truth values.

Proposition 1.2.2 For any statement p, $\neg\neg p$and p have the same value.

p	q	$p \wedge q$	$p \vee q$
T	T	T	T
T	F	F	T
F	T	F	T
F	F	F	F

Table 1.1: Truth table for conjunction and disjunction

1.3 Conjunction and Disjunction

Definition 1.3.1 Let p and q be statements,
We denote by $p \wedge q$ the statement "p and q".
We denote by $p \vee q$ the statement "p or q".

Proposition 1.3.2 Let $P$and Q be statements $(\neg P) \vee (\neg Q)$ and $\neg(P \wedge Q)$ have the same truth value.

1.4 Conditional statements

Definition 1.4.1 Let P and Q be statements, we denote by $P \Rightarrow Q$ the statement (if P then Q).

Remark 1.4.2 It has the same truth value as that of $(\neg P \vee Q)$, only when P is true and Q is false, otherwise it's true .
If one can prove Q is assuming that P is true, then $P \Rightarrow Q$ is true .

Proposition 1.4.3 Let P and Q be statements. If P and $P \Rightarrow Q$ are true, then Q is also true.

Proposition 1.4.4 Let P, Q, R be statements. If $P \Rightarrow Q$ and $Q \Rightarrow R$ are true, then $P \Rightarrow R$ is also true.

Theorem 1.4.5 Let $P$and Q be statements. $P \Rightarrow Q$ and $(\neg Q) \Rightarrow (\neg P)$ have the same truth value.

$(\neg Q) \Rightarrow (\neg P)$ is called the contraposition of $P \Rightarrow Q$, if we prove $(\neg Q) \Rightarrow (\neg P)$, then $P \Rightarrow Q$ is also true.

Example 1.4.6 Prove that , let n be an integer, if n^2 is even, then n is even.

Proof Since $n$is an integer, there exists$k \in \mathbb{Z}$ such that $n = 2k + 1$. Hence $n^2 = 4k^2 + 4k + 1$ is not even. □

1.5 Biconditional statement

Definition 1.5.1 Let P and Q be statements. We denote by $P \Leftrightarrow Q$ the statement

"P if and only if Q"

its true when P and Q have the same truth value, it's false when they have the opposite truth value.

Proposition 1.5.2 Let P and Q be statements. $P \Leftrightarrow Q$ has the same truth value as

$$(P \Rightarrow Q) \wedge (Q \Rightarrow P).$$

Example 1.5.3 Let n be an integer. n is even if and only if n^2 is even.

Definition 1.5.4 Let P and Q be statements.
$Q \Rightarrow P$ is called the converse of $P \Rightarrow Q$.
$\neg P \Rightarrow \neg Q$ is called the inverse of $P \Rightarrow Q$.

Remark 1.5.5 If one proves $P \Rightarrow Q$ and $\neg P \Rightarrow \neg Q$, then $P \Leftrightarrow Q$ is true.

1.6 Proof by Contradiction

Definition 1.6.1 Let P be a statement. If we assume $\neg P$ is true and deduce that a certain statement is both true and false, then we say that a contradiction happens and the assumption $\neg P$ is false. Thus the statement P is true. Such a reasoning is called proof by contradiction.

Example 1.6.2 Prove that the equation $x^2 = 2$ does not have solution in $\mathbb{Q}$.

Proof By contradiction, we assume that $x := \frac{p}{q}$ is a solution, where p and q are integers , which do not have common prime divisor. By $x^2 = 2$ we obtain $p^2 = 2q^2$ So p^2 is even, p is even. Let $p_1 \in \mathbb{Z}$ such that $p = 2p_1$. Then $p^2 = 4p_1^2 = 2q^2$, hence q is even. Therefore 2 is a common prime divisor of p and q, which leads to a contradiction. $\square$

1.7 Exercises

1. Let P and Q be statements. Use truth tables to determine the truth values of the following statements according to the truth values of P and Q:
$$P \wedge \neg P,\ P \vee \neg P,\ (P \vee Q) \Rightarrow (P \wedge Q),\ (P \Rightarrow Q) \Rightarrow (Q \Rightarrow P)$$

2. Let P and Q be statements.
   (a) Show that $P \Rightarrow (Q \wedge \neg Q)$ has the same truth value as $\neg P$.
   (b) Show that $(P \wedge \neg Q) \Rightarrow Q$ has the same truth value as $P \Rightarrow Q$.

3. Consider the following statements:
$$P := \text{“Little Bear is happy”},$$
$$Q := \text{“Little Bear has done her math homework”},$$
$$R := \text{“Little Rabbit is happy”}.$$
Express the following statements using P, Q, and R, along with logical connectives:
   (a) If Little Bear is happy and has done her math homework, then Little Rabbit is happy.
   (b) If Little Bear has done her math homework, then she is happy.
   (c) Little Bear is happy only if she has done her math homework.

4. Does the following reasoning hold? Justify your answer.
   - It is known that Little Bear is both smart and lazy, or Little Bear is not smart.

- It is also known that Little Bear is smart.
- Therefore, Little Bear is lazy.

5. Does the following reasoning hold? Justify your answer.

   - It is known that at least one of the lion or the tiger is guilty.
   - It is also known that either the lion is lying or the tiger is innocent.
   - Therefore, the lion is either lying or guilty.

6. An explorer arrives at a cave with three closed doors, numbered 1, 2, and 3. Exactly one door hides treasure, while the other two conceal deadly traps.

   - Door 1 states: "*The treasure is not here*";
   - Door 2 states: "*The treasure is not here*";
   - Door 3 states: "*The treasure is behind Door 2*".

   Only one of these statements is true. Which door should the explorer open to find the treasure?

7. The Kingdom of Truth sent an envoy to the capital of the Kingdom of Lies. Upon entering the border, the envoy encountered a fork with three paths: dirt, stone, and concrete. Each path had a signpost:

   - The concrete path's sign: "*This path leads to the capital, and if the dirt path leads to the capital, then the stone path also does.*"
   - The stone path's sign: "*Neither the concrete nor the dirt path leads to the capital.*"
   - The dirt path's sign: "*The concrete path leads to the capital, but the stone path does not.*"

   All signposts lie. Which path should the envoy take?

8. Let a and b be real numbers. Prove that, if $a \neq -1$ and $b \neq -1$, then $ab + a + b \neq -1$.

9. Let a, b, and c be positive real numbers such that $abc > 1$ and

   $$a + b + c < \frac{1}{a} + \frac{1}{b} + \frac{1}{c}.$$

   Prove the following:

   (a) None of a, b, or c equals 1.

(b) At least one of a, b, or c is greater than 1.

(c) At least one of a, b, or c is less than 1.

10. Let $a \neq 0$ and b be real numbers. For real numbers x and y, prove that if $x \neq y$, then $ax + b \neq ay + b$.

11. Let $n \geq 2$ be an integer. Prove that if n is composite, then there exists a prime number p dividing n such that $p \leq \sqrt{n}$.

12. Let n be an integer. Prove that either 4 divides n^2 or 4 divides $n^2 - 1$.

13. Let n be an integer. Prove that 12 divides $n^2(n^2 - 1)$.

14. Prove that any integer divisible by 4 can be written as the difference of two perfect squares.

15. Let x and y be non-zero integers. Prove that $x^2 - y^2 \neq 1$.

16. A plane has 300 seats and is fully booked. The first passenger ignores their assigned seat and chooses randomly. Subsequent passengers take their assigned seat if available; otherwise, they choose randomly. What is the probability that the last passenger sits in their assigned seat?

17. Little Bear, Little Goat, and Little Rabbit are all wearing hats. A parrot prepared four red feathers and four blue feathers to decorate their hats. The parrot selected two feathers for each hat-wearing animal to place on their hats. Each animal cannot see the feathers on their own hat but can see the feathers on the other animals' hats. Here is their conversation:

    - Little Bear: *I don't know what color the feathers on my hat are, but I know the other animals also don't know what color the feathers on their hats are.*
    - Little Goat: *Haha, now even without looking at Little Bear's hat, I know what color the feathers on my hat are.*
    - Little Rabbit: *Now I know what color the feathers on my hat are.*
    - Little Bear: *Hmm, now I also know what color the feathers on my hat are.*

    Question: What color are the feathers on Little Goat's hat?

18. The Sphinx tells the truth on one fixed weekday and lies on the other six. Cleopatra visits The Sphinx for three consecutive days:

    - Day 1: The Sphinx declared, "*I lie on Monday and Tuesday.*"

- Day 2: The Sphinx declared, "*Today is either Thursday, or Saturday, or Sunday.*"
- Day 3: The Sphinx declared, "*I lie on Wednesday and Friday.*"

On which day does the Sphinx tell the truth? On which days of the week did Cleopatra visit the Sphinx?

Chapter 2

Set Theory

2.1 Roster Notation

Definition 2.1.1
(1) We call a **set** a certain collection of distinct objects.
(2) An object in a collection considered as a set is called **element** of it .
(3) Two sets A and B are said to be **equal** if they have the same elements.We denoted by $A = B$ the statement “A and B are equal”.
(4) If A is a set and x is an object, $x \in A$ denotes x is an element of A (reads x belongs to A), $x \notin A$ denotes “x is NOT an element of A”.

`Notation Roster method: to be continue...`

Example 2.1.2 $\{1, 2, 3\} = \{3, 2, 1\} = \{1, 1, 2, 3\}$.

More generally, if I is a set, and for any $i \in I$, we fix an x_i, then the set of all x_i is noted as

$$\{x_i \mid i \in I\}.$$

Example 2.1.3

$$\{2k + 1 \mid k \in \mathbb{Z}\}.$$

2.2 Set-builder Notation

Definition 2.2.1 Let A be a set. If for any $x \in A$ we fix a statement $P(x)$, then we say that $P(\cdot)$ is a **condition** on A.

Example 2.2.2 "n is even" is a condition on $\mathbb{N}$, "$x > 2$" is a condition on $\mathbb{R}$.

Definition 2.2.3 Let A be a set and $P(\cdot)$ be a condition on A. If $x \in A$ is such that $P(x)$ is true, then we say that x satisfies the condition $P(\cdot)$. We noted by

$$\{x \in A \mid P(x)\}$$

the set of $x \in A$ that satisfies the condition $P(\cdot)$.

Example 2.2.4 $\{x \in \mathbb{R} | x > 2\}$ denotes the set of real numbers that are $x > 2$.

sometimes we combine the two methods of representation.

2.3 Subsets and Set Difference

Definition 2.3.1 Let A and B be sets. If any element of A is an element of B, we say that A is a subset of B, denoted as $A \subseteq B$ or $B \supseteq A$.

Example 2.3.2

- We denote by $\varnothing$ the set that does not have any element. We consider it as a subset of any set.
- Let A be a set, then $A \subseteq A$.

Definition 2.3.3 Let A be a set, we denote by $\mathscr{P}(A)$ the set of all subset of A, called the power set of A.

Example 2.3.4 $\mathscr{P}(\varnothing) = \{\varnothing\}$, $\mathscr{P}(\mathscr{P}(\varnothing)) = \{\varnothing, \{\varnothing\}\}$.

Definition 2.3.5 Let A and B be sets. We denote by $B\backslash A$ the set

$$\{x \in B \mid x \notin A\}.$$

This is a subset of B called the **set difference of B and A**.
If in condition $A \subseteq B$, we say that $B\backslash A$ is the complement of A inside B.

Example 2.3.6 If A is a set, $P(\cdot)$ is a condition on A, then

$$\{x \in A \mid \neg P(x)\} = A\backslash\{x \in A \mid P(x)\}.$$

Proposition 2.3.7 Let A and B be sets. Then

$$B\backslash A = \varnothing \Leftrightarrow B \subseteq A.$$

If in condition A is the subset of B, then

$$B\backslash A = \varnothing \Leftrightarrow A = B.$$

2.4 Quantifiers

Definition 2.4.1 Let A be a set and $P(\cdot)$ be a condition on A .We denote by "$\forall x \in A, P(x)$" the statement $\{x \in A \mid P(x)\} = A$.
"$\exists x \in A, P(x)$" denotes$\{x \in A \mid P(x)\} \neq \varnothing$.

Example 2.4.2 $\forall x \in \varnothing, P(x)$ is true ; $\exists x \in \varnothing, P(x)$ is false.

Theorem 2.4.3 Let A be a set and $P(\cdot)$ be a condition on A
(1) $\exists x \in A, \neg P(x)$ and $\forall x \in A, P(x)$ have opposite truth values.
(2) $\forall x \in A, \neg P(x)$ and $\exists x \in A, P(x)$ have opposite truth value.

2.5 Sufficient and Necessary Condition

Definition 2.5.1 Let A be a set and $P(\cdot)$ and $Q(\cdot)$ be conditions on A. If

$$\{x \in A \mid P(x)\} \subseteq \{x \in A \mid Q(x)\},$$

we say that $P(\cdot)$ is a **sufficient condition** of $Q(\cdot)$ and $Q(\cdot)$ is a **necessary condition** of $P(\cdot)$. If $\{x \in A \mid P(x)\} = \{x \in A \mid Q(x)\}$, we say that $P(\cdot)$ and $Q(\cdot)$ are equivalent.

Proposition 2.5.2 Let A be a set, $P(\cdot)$ and $Q(\cdot)$ be conditions on A.
(1) $P(\cdot)$ is a sufficient condition of $Q(\cdot)$ if and only if $\forall x \in A, P(x) \Rightarrow Q(x)$.
(2) $P(\cdot)$ is a necessary condition of $Q(\cdot)$ if and only if $\forall x \in A, Q(x) \Rightarrow P(x)$.
(3) $P(\cdot)$ and $Q(\cdot)$ are equivalent if and only if $\forall x \in A, P(x) \Leftrightarrow Q(x)$.

Proof

$$\begin{aligned}\varnothing &= \{x \in A \mid P(x)\} \backslash \{x \in A \mid Q(x)\} \\ &= \{x \in A \mid P(x) \wedge (\neg Q(x))\} \\ &= A \backslash \{x \in A \mid (\neg P(x)) \vee Q(x)\} \\ &= A \backslash \{x \in A \mid P(x) \Rightarrow Q(x)\}.\end{aligned}$$

□

Russell's paradox leads to: $P(A) := A \notin A$. The collection of all sets should not be considered as a set.

2.6 Union

Definition 2.6.1 Let I be a set, and for any $i \in I$. Let A_i be a set, we say that $(A_i)_{i \in I}$ is a family of sets parametrized by I. We denote by $\cup_{i \in I} A_i$ the set of all elements of all A_i. It is also called the **union** of the sets A_i, $i \in I$. By definition, a mathematical object x belongs to $\cup_{i \in I} A_i$ if and only if

$$\exists i \in I, x \in A_i.$$

Proposition 2.6.2 $\bigcup\limits_{i \in I} A_i \subseteq B$ if and only if

$$\forall i \in I, A_i \subseteq B.$$

Corollary 2.6.3 Let $P_i(\cdot)$ be a condition on B, then

$$\{x \in B \mid \exists i \in I, P_i(x)\} = \bigcup_{i\in I}\{x \in B \mid P_i(x)\}.$$

Proposition 2.6.4

$$\left(\bigcup_{i\in I} A_i\right) \backslash B = \bigcup_{i\in I}\left(A_i \backslash B\right).$$

2.7 Intersection

Definition 2.7.1 Let I be a **non-empty** set and $(A_i)_{i\in I}$ be a family os sets parametrized by I. We denote by $\bigcap\limits_{i\in I} A_i$ the set of all common elements of A_i, $i \in I$. This set is called the **intersection** of $A_i, i \in I$. Note that, if i_0 is an arbitrary element of I, the set-builder notation ensure that

$$\{x \in A_{i_0} \mid \forall i \in I, x \in A_i\}$$

is a set. This set is the intersection of $(A_i)_{i\in I}$.
By definition, an mathematical object x belongs to $\cap_{i\in I} A_i$ if and only if

$$\forall i \in I, x \in A_i.$$

Remark 2.7.2 In set theory, it does not make sense to consider the intersection of an empty family of sets. In fact, if such an intersection exists as a sets, for any mathematical object x, since the statement

$$\forall i \in \varnothing, x \in A_i$$

is true, we obtain that x belongs to $\cap_{i\in\varnothing} A_i$. By Russell's paradox, this is impossible.

Proposition 2.7.3 Let I be a non-empty set and $(A_i)_{i\in I}$ be a set parametrised by I. Let B be a set. Then $B \subseteq \bigcap\limits_{i\in I} A_i$ if and only if

$$\forall\, i \in I,\ B \subseteq A_i.$$

Proof Let $A = \bigcap_{i\in I} A_i$.
Suppose that $B \subseteq A$. For any $x \in B$, one has $x \in A$, and hence

$$\forall\, i \in I,\ x \in A_i.$$

Therefore, for any $i \in I$, B is contained in A_i.
Suppose that, for any $i \in I$, $B \subseteq A_i$. Then, for any $x \in B$ and any $i \in I$, one has $x \in A_i$. Hence, for any $x \in B$, one has $x \in A$. Therefore, $B \subseteq A$. □

Corollary 2.7.4 Let B be a set, I be a non-empty set. For any $i \in I$, let $P_i(\cdot)$ be a condition on B. Then

$$\{x \in B \mid \forall\, i \in I,\ P_i(x)\} = \bigcap_{i\in I} \{x \in B \mid P_i(x)\}.$$

Proof Let

$$A := \{x \in B \mid \forall\, i \in I,\ P_i(x)\}.$$

For any $i \in I$, let

$$A_i := \{x \in B \mid P_i(x)\}.$$

For any $x \in A$ and any $i \in I$, $P_i(x)$ is true. Hence $A \subseteq A_i$. By Proposition 2.7.3 we obtain

$$A \subseteq \bigcap_{i\in I} A_i.$$

Conversely, if $x \in \bigcap_{i\in I} A_i$, then for any $i \in I$, one has $x \in A_i$. Hence $x \in B$, and for any $i \in I$, $P_i(x)$ is true. Thus $x \in A$. □

Proposition 2.7.5 Let B be a set, $(A_i)_{i\in I}$ be a family of sets. The following equality holds

$$\left(\bigcap_{i\in I} A_i\right) \setminus B = \bigcap_{i\in I} (A_i \setminus B).$$

Proof Let $A := \bigcap_{i\in I} A_i$. For any $i \in I$, one has $A \subseteq A_i$. Hence

$$A \setminus B = \{x \in A \mid x \notin B\} \subseteq \{x \in A_i \mid x \notin B\}.$$

By Proposition 2.7.3 we get

$$A \setminus B \subseteq \bigcap_{i\in I} (A_i \setminus B).$$

Conversely, if $x \in \bigcap_{i\in I}(A_i \setminus B)$, then, for any $i \in I$, one has $x \in A_i \setminus B$, namely $x \in A_i$ and $x \notin B$. Thus $x \in \bigcap_{i\in I} A_i$ and $x \notin B$. Therefore $x \in A \setminus B$. □

Proposition 2.7.6 Let I be a set and $(A_i)_{i\in I}$ be a family of sets parametrised by I. For any set B, the following statements hold.

1. $B \cap (\bigcup_{i\in I} A_i) = \bigcup_{i\in I}(B \cap A_i)$.
2. If $I \neq \varnothing$, $B \cup (\bigcap_{i\in I} A_i) \subseteq \bigcap_{i\in I}(B \cup A_i)$,
3. If $I \neq \varnothing$, $B \setminus \bigcup_{i\in I} A_i = \bigcap_{i\in I}(B \setminus A_i)$,
4. If $I \neq \varnothing$, $B \setminus \bigcap_{i\in I} A_i = \bigcup_{i\in I}(B \setminus A_i)$.

Proof 1. By Corollary 2.7.4 we obtain

$$\begin{aligned} B \cap \Big(\bigcup_{i\in I} A_i\Big) &= \{x \in B \mid \exists\, i \in I,\ x \in A_i\} \\ &= \bigcup_{i\in I}\{x \in B \mid x \in A_i\} = \bigcup_{i\in I}(B \cap A_i). \end{aligned}$$

2. Let $A := \bigcap_{i\in I} A_i$. By definition, for any $i \in I$, one has $A \subseteq A_i$ and hence $B \cup A \subseteq B \cup A_i$. Thus, by Proposition 2.7.3 we obtain

$$B \cup \Big(\bigcap_{i\in I} A_i\Big) \subseteq \bigcap_{i\in I}(B \cup A_i).$$

Conversely, let $x \in \bigcap_{i\in I}(B \cup A_i)$. For any $i \in I$, one has $x \in B \cup A_i$. If $x \in B$, then $x \in B \cup (\bigcap_{i\in I} A_i)$; otherwise one has

$$\forall\, i \in I,\ x \in A_i,$$

and we still get $x \in B \cup (\bigcap_{i\in I} A_i)$.

3. By Theorem2.4.3

$$\begin{aligned} B \setminus \bigcup_{i\in I} A_i &= \{x \in B \mid \neg(\exists\, i \in I,\ x \in A_i)\} \\ &= \{x \in B \mid \forall\, i \in I,\ x \notin A_i\}. \end{aligned}$$

By Corollary 2.7.4 this is equal to

$$\bigcap_{i\in I}\{x \in B \mid x \notin A_i\} = \bigcap_{i\in I}(B \setminus A_i).$$

4. By Theorem2.4.3

$$
\begin{aligned}
B \setminus \bigcap_{i \in I} A_i &= \{x \in B \mid \neg(\forall\, i \in I,\ x \in A_i)\} \\
&= \{x \in B \mid \exists\, i \in I,\ x \notin A_i\}.
\end{aligned}
$$

By Corollary 2.6.3 this is equal to

$$\bigcup_{i \in I} \{x \in B \mid x \notin A_i\} = \bigcup_{i \in I} (B \setminus A_i).$$

□

2.8 Cartesian Product

Definition 2.8.1 Let A and B be sets. We denote by $A \times B$ the following set of ordered pairs

$$\{(x, y) \mid x \in A,\ y \in B\},$$

and call it the **Cartesian product** of sets A and B.
More generally, if n is a positive integer and $A_1, \ldots, A_n$ be sets, we denote by

$$A_1 \times \cdots \times A_n$$

the set of all n-tuples $(x_1, \ldots, x_n)$, where $x_1 \in A_1, \ldots, x_n \in A_n$.

The following proposition shows ordered pairs can be realized through set-theoretic constructions.

Proposition 2.8.2 Let x, y, x', and y' be mathematical objects. Then

$$\{\{x\}, \{x, y\}\} = \{\{x'\}, \{x', y'\}\}$$

if and only if $x = x'$ and $y = y'$.

Proof If $x = x'$ and $y = y'$, then the equality

$$\{\{x\}, \{x, y\}\} = \{\{x'\}, \{x', y'\}\}$$

certainly holds.

Conversely, suppose the equality

$$\{\{x\},\{x,y\}\} = \{\{x'\},\{x',y'\}\}$$

holds. If $x \neq x'$, then $\{x\} \neq \{x'\}$, so $\{x\} = \{x', y'\}$. This still implies $x = x'$, leading to a contradiction. Therefore, $x = x'$ must hold.
Now, assume $y \neq y'$. Then $\{x, y\} \neq \{x', y'\}$, unless $y = x'$ and $x = y'$. Since $x = x'$, this would imply $y = y'$, which is a contradiction. Thus, $\{x, y\} = \{x'\}$ and $\{x', y'\} = \{x\}$. This again leads to $y = x'$ and $x = y'$, resulting in a contradiction. Hence, $y = y'$ must hold. $\square$

Chapter 3

Correspondence

3.1 Correspondence and its Inverse

Definition 3.1.1 We call a **correspondence** any triplet of the form

$$f = (\mathscr{D}_f, \mathscr{A}_f, \Gamma_f)$$

where $\mathscr{D}_f, \mathscr{A}_f$ are two sets, called respectively the **departure set** and the **arrival set** of f and Γ_f is a subset of $\mathscr{D}_f \times \mathscr{A}_f$, called the **graph** of f.
If X, Y are two sets and f is a correspondence of the form (X, Y, Γ_f), we say that f is a correspondence from X to Y.

Definition 3.1.2 Let f be a correspondence.We denote by f^{-1} the correspondence defined as follows:

$$\mathscr{D}_f^{-1} := \mathscr{A}_f, \mathscr{A}_f^{-1} := \mathscr{D}_f,$$

$$\Gamma_{f^{-1}} := \{(y, x) \in \mathscr{D}_f \times \mathscr{A}_f \mid (x, y) \in \Gamma_f\}.$$

The correspondence f^{-1} is called the **inverse correspondence** of f. Clearly one has

$$(f^{-1})^{-1} = f,$$

namely f is the inverse correspondence of f^{-1}.

3.2 Illustration of a Correspondence

3.3 Image and Preimage

Definition 3.3.1 Let X, Y be sets , and f be a correspondence from X to Y. If (x, y) is an element of Γ_f, we say that x is a **preimage** of y under f, and y is an **image** of x under f.
If A is a set , we denote by $f(A)$ the set :

$$\{y \in \mathscr{A}_f \mid \exists x \in A, (x, y) \in \Gamma_f\},$$

called the image of A by the correspondence f.
If B is a set , the set $f^{-1}(B)$ is called the **preimage of B by the correspondence** f. Note that it is by definition the image of B by the inverse correspondence f^{-1}.

Definition 3.3.2 Let f be correspondence. The set $f(\mathscr{D}_f)$ is called the **range** of f, denoted as $\mathrm{Im}(f)$. The set $f^{-1}(\mathscr{A}_f)$ is called the **domain of definition** of f, denoted as $\mathrm{Dom}(f)$. Note that the domain of definition of a correspondence f is the projection of the graph Γ_f to the arrival set $\mathscr{A}_f$.

For any sets A and B,

$$f(A) \subseteq \mathrm{Im}(f), f^{-1}(B) \subseteq \mathrm{Dom}(f),$$

$$\mathrm{Dom}(f) = \mathrm{Im}(f^{-1}), \mathrm{Im}(f) = \mathrm{Dom}(f^{-1}).$$

Proposition 3.3.3 Let f be a correspondence.
(1) If A and A' are two sets such that $A' \subseteq A$, then one has $f(A') \subseteq f(A)$.
(2) If B and B' are two sets such that $B' \subseteq B$, then one has $f^{-1}(B') \subseteq f^{-1}(B)$.

Proof

$$\begin{aligned} f(B') &= \{y \in \mathrm{Im}(f) \mid \exists x \in B', (x, y) \in \Gamma_f\} \\ &\subseteq \{y \in \mathrm{Im}(f) \mid \exists x \in B, (x, y) \in \Gamma_f\} \\ &= f(B). \end{aligned}$$

□

Proposition 3.3.4 Let f be a correspondence. The following equalities hold:

$$\mathrm{Im}(f) = f(\mathrm{Dom}(f)), \mathrm{Dom}(f) = f^{-1}(\mathrm{Im}(f)).$$

Proof Since $\mathrm{Dom}(f) \subseteq \mathscr{D}_f$, by proposition 3.3.3, one has

$$f(\mathrm{Dom}(f)) \subseteq f(\mathscr{D}_f) = \mathrm{Im}(f).$$

Let y be an element of $\mathrm{Im}(f)$, there exist $x \in \mathscr{D}_f$ such that $(x, y) \in \Gamma_f$. By definition, one has $x \in \mathrm{Dom}(f)$ and hence $y \in f(\mathrm{Dom}(f)), \mathrm{Im}(f) \subseteq f(\mathrm{Dom}(f))$. Therefore the equality $\mathrm{Im}(f) = f(\mathrm{Dom}(f))$ is true. Applying this equality to f^{-1}, we obtain the second equality. □

Proposition 3.3.5 Let f be a correspondence, A be a set and y be an mathematical object. Then y belongs to $f(A)$ if and only if $A \cap f^{-1}(\{y\}) \neq \varnothing$.

Proposition 3.3.6 Let f be a correspondence, I be a set and $(A_i)_{i \in I}$ be a family of sets parametrised by I. Then

$$f\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} f(A_i).$$

Moreover, if I is not empty, then

$$f\left(\bigcap_{i \in I} A_i\right) \subseteq \bigcap_{i \in I} f(A_i).$$

Proof

$$\begin{aligned} f\left(\bigcup_{i \in I} A_i\right) &= \left\{ y \in Y \mid \left(\bigcup_{i \in I} A_i\right) \cap f^{-1}(\{y\}) \neq \varnothing \right\} \\ &= \left\{ y \in Y \mid \bigcup_{i \in I} \left(A_i \cap f^{-1}(\{y\})\right) \neq \varnothing \right\} \\ &= \left\{ y \in Y \mid \exists i \in I, A_i \cap f^{-1}(\{y\}) \neq \varnothing \right\} = \bigcup_{i \in I} f(A_i). \end{aligned}$$

$$\begin{aligned} f\left(\bigcap_{i\in I} A_i\right) &= \left\{y \in Y \mid \left(\bigcap_{i\in I} A_i\right) \cap f^{-1}(\{y\}) \neq \varnothing\right\} \\ &= \left\{y \in Y \mid \bigcap_{i\in I} \left(A_i \cap f^{-1}(\{y\})\right) \neq \varnothing\right\} \\ &\subseteq \left\{y \in Y \mid \forall i \in I, A_i \cap f^{-1}(\{y\}) \neq \varnothing\right\} \\ &= \bigcap_{i\in I} f(A_i). \end{aligned}$$

□

3.4 Composition

Definition 3.4.1 Let f and g be correspondences. We define the **composite** of g and f as the correspondence $g \circ f$ from $\mathscr{D}_f$ to $\mathscr{A}_g$ whose graph $\Gamma_{g\circ f}$ is composed of the element (x, z) of $\mathscr{D}_f \times \mathscr{A}_g$ such that there exists some objet y satisfying $(x, y) \in \Gamma_f$ and $(y, z) \in \Gamma_g$. In other words,

$$\Gamma_{g\circ f} = \{(x, z) \in \mathscr{D}_f \times \mathscr{A}_g \mid \exists y \in \mathscr{A}_f \cap \mathscr{D}_g, (x, y) \in \Gamma_f \wedge (y, z) \in \Gamma_g\}.$$

Proposition 3.4.2 Let f and g be correspondences. The following equality holds:

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}. \tag{3.4.1}$$

Proposition 3.4.3 Let f and g be correspondences.The following equality holds:

$$h \circ (g \circ f) = (h \circ g) \circ f. \tag{3.4.2}$$

Proposition 3.4.4 Let X and Y be sets , f be a correspondence from X to Y.Then the following equalities hold:

$$f \circ \mathrm{Id}_X = f = \mathrm{Id}_Y \circ f.$$

Propositions above can be proved by definition.

Proposition 3.4.5 Let f and g be correspondences.For any set A , one has

$$(g \circ f)(A) = g(f(A)).$$

In particular,

$$\mathrm{Im}(g \circ f) = g(\mathrm{Im}(f)) \subseteq \mathrm{Im}(g).$$

If in addition $\mathrm{Dom}(g) \subseteq \mathrm{Im}(f)$, then the equality $\mathrm{Im}(g \circ f) = \mathrm{Im}(g)$ holds.

Proof By definition,

$$\begin{aligned}(g \circ f)(A) &= \{z \in \mathscr{A}_g \mid \exists x \in A, (x, z) \in \Gamma_{g \circ f}\} \\ &= \{z \in \mathscr{A}_g \mid \exists x \in A, \exists y \in \mathscr{A}_f, (x, y) \in \Gamma_f, (y, z) \in \Gamma_g\} \\ &= \{z \in \mathscr{A}_g \mid \exists y \in f(A), (y, z) \in \Gamma_g\} = g(f(A)).\end{aligned}$$

Applying this equality to the case where $A = \mathscr{D}_f$, we obtain

$$\mathrm{Im}(g \circ f) = (g \circ f)(\mathscr{D}_f) = g(f(\mathscr{D}_f)) = g(\mathrm{Im}(f)) \subseteq \mathrm{Im}(g).$$

In the case where $\mathrm{Dom}(g) \subseteq \mathrm{Im}(f)$, by proposition 3.3.3 and 3.3.4 we obtain

$$\mathrm{Im}(g) = g(\mathrm{Dom}(g)) \subseteq g(\mathrm{Im}(f)) = \mathrm{Im}(g \circ f).$$

□

3.5 Surjectivity

Definition 3.5.1 Let f be a correspondence. If $\mathscr{A}_f = \mathrm{Im}(f)$, we say that f is **surjective**. If f^{-1} is surjective , or equivalently $\mathrm{Dom}(f) = \mathscr{D}_f$, we say that f is a **multivalued mapping**.

Remark 3.5.2 multivalued mapping is not always a mapping

Proposition 3.5.3 Let f be a correspondence.Assume that f is surjective. Then , for any subset B of $\mathscr{A}_f$, one has $B \subseteq f(f^{-1}(B))$.

Proof Let y be an element of B. Since f is surjective there exists $x \in \mathscr{D}_f$ such that $(x, y) \in \Gamma_f$ .Therefore, $x \in f^{-1}(B)$ and hence $y \in f(f^{-1}(B))$ □

Proposition 3.5.4 Let f and g be correspondences.
(1) If $g \circ f$ is surjective, so is g.

(2) If $g \circ f$ is multivalued mapping, so is f.

Proof One has

$$\mathrm{Im}(g \circ f) \subseteq \mathrm{Im}(g) \subseteq \mathscr{A}_g = \mathscr{A}_{g \circ f}.$$

If $g \circ f$ is surjective, namely $\mathrm{Im}(g \circ f) = \mathscr{A}_{g \circ f}$, then we deduce $\mathrm{Im}(g) = \mathscr{A}_g$, namely g is surjective. □

Proposition 3.5.5 Let f and g be correspondences.
(1) If g is surjective and $\mathrm{Dom}(g) \subseteq \mathrm{Im}(f)$, then $g \circ f$ is also surjective.
(2) If f is a multivalued mapping and $\mathrm{Im}(f) \subseteq \mathrm{Dom}(g)$, then $g \circ f$ is a multivalued mapping.

Proof (1) Since $\mathrm{Dom}(g) \subseteq \mathrm{Im}(f)$, by proposition3.4.5, we obtain

$$\mathrm{Im}(g \circ f) = \mathrm{g}.$$

Since g is surjective,

$$\mathrm{Im}(g) = \mathscr{A}_g = \mathscr{A}_{g \circ f}.$$

Hence $g \circ f$ is also surjective.
Applying (1) to g^{-1} and f^{-1} , we obtain (2). □

3.6 injectivity

Definition 3.6.1 Let f be a correspondence. If each element of $\mathscr{D}_f$ has at most one image under f , we say that f is a **function**.If f^{-1} is a function, we say that f is **injective** .

Notation 3.6.2 Functions form a special case of correspondences.The definition feature of functions is that corresponding to each element in the domain of definition, is a unique element in the arrival set of function.
Let f be a function, and let $x \in \mathrm{Dom}(f)$. We denote the unique image of x under f as $f(x)$, and we say that f sends $x \in \mathrm{Dom}(f)$ to $f(x)$ or $f(x)$ is the **value** of f at x .we can also use the notation:

$$x \mapsto f(x)$$

to indicate the correspondence of x to its image under f.

Proposition 3.6.3 Let f be a correspondence.
(1) Assume that f is injective.For any set A one has $f^{-1}(f(A)) \subseteq A$.
(2) Assume that f is a function. For any set B on has $f(f^{-1}(B)) \subseteq B$.

Proof Let x be an element of $f^{-1}(f(A))$, By definition, there exists $y \in f(A)$ such that $(x, y) \in \Gamma_f$. Since $y \in f(A)$ there exist $x' \in A$ such that $(x', y) \in \Gamma_f$.Since y admits at most one preimage, we obtain $x' = x$. Hence $x \in A$.
Applying (1) to f^{-1} we obtain (2). □

Proposition 3.6.4 Let f and g be correspondences.
(1) If f and g are functions, so is $g \circ f$. Moreover, for any $x \in \mathrm{Dom}(g \circ f)$, one has $(g \circ f)(x) = g(f(x))$.
(2) If f and g are injective, so is $g \circ f$.

Proof Let x be an element of $\mathrm{Dom}(g \circ f)$. Assume that z and z' are images of x under $g \circ f$. Let y and y' be such that

$$(x, y) \in \Gamma_f, \quad (y, z) \in \Gamma_g, \quad (x, y') \in \Gamma_f, \quad (y', z') \in \Gamma_g.$$

Since f is a function, one has $y = y' = f(x)$. Since g is a function, we deduce that $z = z' = g(f(x))$. Therefore $g \circ f$ is a function, and the equality $(g \circ f)(x) = g(f(x))$ holds for any $x \in \mathrm{Dom}(g \circ f)$.
Applying (1) to g^{-1} and f^{-1}, we obtain (2). □

Proposition 3.6.5 Let f and g be correspondences.
(1) If $g \circ f$ is injective and $\mathrm{Im}(f) \subseteq \mathrm{Dom}(g)$, then f is also injective.
(2) If $g \circ f$ is a function and $\mathrm{Dom}(g) \subseteq \mathrm{Im}(f)$, then g is also a function.

Proof
(1) Let y be an element of the image of f. Let x and x' be preimages of y under f. Since $\mathrm{Im}(f) \subseteq \mathrm{Dom}(g)$, one has $y \in \mathrm{Dom}(g)$. Hence there exists $z \in \mathscr{A}_g$ such that $(y, z) \in \Gamma_g$. We then deduce that (x, z) and (x', z) are elements of $\Gamma_{g \circ f}$. Since $g \circ f$ is injective, we obtain $x = x'$. Therefore, f is injective.
Applying (1) to g^{-1} and f^{-1}, we obtain (2). □

Proposition 3.6.6 Let f be a correspondence, and I be a non-empty set.
(1) Suppose that f is a function. For any family $(B_i)_{i \in I}$ of sets parametrised by

I, one has

$$f^{-1}\left(\bigcap_{i\in I} B_i\right) = \bigcap_{i\in I} f^{-1}(B_i).$$

(2) Suppose that f is injective. For any family $(A_i)_{i\in I}$ of sets parametrised by I, one has

$$f\left(\bigcap_{i\in I} A_i\right) = \bigcap_{i\in I} f(A_i).$$

Proof
(1) Let x be an element of $\bigcap_{i\in I} f^{-1}(B_i)$. For any $i \in I$, one has $f(x) \in B_i$. Hence $x \in f^{-1}(\bigcap_{i\in I} B_i)$. Therefore we obtain

$$f^{-1}\left(\bigcap_{i\in I} B_i\right) \supseteq \bigcap_{i\in I} f^{-1}(B_i).$$

Combining with (2) of proposition3.3.6 , we obtain the equality

$$f^{-1}\left(\bigcap_{i\in I} B_i\right) = \bigcap_{i\in I} f^{-1}(B_i).$$

Applying (1) to f^{-1}, we obtain (2). □

3.7 Mapping

Definition 3.7.1 A correspondence f is said to be a **mapping** if any element of $\mathscr{D}_f$ has a unique image, or equivalently, f is a function and $\mathscr{D}_f = \mathrm{Dom}(f)$. Note that f is a mapping if and only if f^{-1} is both injective and surjective.

Notation 3.7.2 Let X and Y be sets. We denote by Y^X the set of all mappings from X to Y. An element $u \in Y^X$ is often written in the form of a family of elements of Y parametrised by X as follows

$$(u(x))_{x\in X}.$$

In the case where $X = \{1, \ldots, n\}$, where n is a positive integer, the set $Y^{\{1,\ldots,n\}}$ is also denoted as Y^n. An element u of Y^n is often written as

$$(u(1), \ldots, u(n)).$$

Example 3.7.3

1. Let X be a set. The identity correspondence Id_X is a mapping. It is also called the **identity mapping** of X.
2. Let X and Y be sets and y be an element of Y. The mapping from X to Y sending any $x \in X$ to y is called the **constant mapping with value** y.
3. Let X be a set and $A \subseteq X$, we define $\mathbb{1}_A : X \to \mathbb{R}$

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

   It is called **indicator function**

Remark 3.7.4 Let $f : X \to Y$ be a mapping, I be a set.

1. By (1) of Proposition3.3.6, for any family of sets $(A_i)_{i\in I}$, one has

$$f\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} f(A_i).$$

   By (2) of Proposition 3.3.6, for any family of sets $(B_i)_{i\in I}$, one has

$$f^{-1}\left(\bigcup_{i\in I} B_i\right) = \bigcup_{i\in I} f^{-1}(B_i).$$

2. Assume that I is not empty. By (1) of Proposition 3.3.6, for any family of sets $(A_i)_{i\in I}$, one has

$$f\left(\bigcap_{i\in I} A_i\right) \subseteq \bigcap_{i\in I} f(A_i).$$

   By (1) of Proposition 3.6.6, for any family of sets $(B_i)_{i\in I}$, one has

$$f^{-1}\left(\bigcap_{i\in I} B_i\right) = \bigcap_{i\in I} f^{-1}(B_i).$$

3. By (2) of Proposition 3.6.3, for any set B, one has $f(f^{-1}(B)) \subseteq B$. Since f is a function and f^{-1} is injective, by (1) of Proposition 3.6.3 and (2) of Proposition 3.5.3, for any subset A of X one has $f^{-1}(f(A)) = A$.

Proposition 3.7.5 Let f and g be mappings. Suppose that $\mathrm{Im}(f) \subseteq \mathscr{D}_g$. Then $g \circ f$ is also a mapping. Moreover, for any $x \in \mathscr{D}_f = \mathscr{D}_{g\circ f}$ one has

$$(g \circ f)(x) = g(f(x)).$$

Proof Note that $\mathscr{D}_g = \mathrm{Dom}(g)$ since g is a mapping. Hence the statement is a direct consequence of Propositions 3.6.4 and 3.5.5

□

Remark 3.7.6 Let $f : X \to Y$ and $g : Y \to Z$ be mappings.

1. By Proposition 3.5.5, if f and g are both surjective, so is $g \circ f$. By Proposition 3.5.4, if $g \circ f$ is surjective, so is g.
2. By Proposition 3.6.4, if f and g are both injective, so is $g \circ f$. By Proposition 3.6.5, if $g \circ f$ is injective, so is f.

3.8 Bijection

Definition 3.8.1 Let f be a mapping, that is, a correspondence such that f^{-1} is injective and surjective. If f is injective and surjective, we say that f is a **bijection**, or a **one-to-one correspondence**. Note that a correspondence is a bijection if and only if its inverse is a bijection.

Proposition 3.8.2 Let X and Y be sets, f be a correspondence from X to Y. If f is a bijection, then $f^{-1} \circ f = \mathrm{Id}_X$ and $f \circ f^{-1} = \mathrm{Id}_Y$. Conversely, if there exists a correspondence g such that $g \circ f = \mathrm{Id}_X$ and $f \circ g = \mathrm{Id}_Y$, then f is a bijection and $g = f^{-1}$.

Proof If f is a bijection, then f and f^{-1} are both mappings. By Proposition 3.7.5, one has

$$\forall x \in X, \quad (f^{-1} \circ f)(x) = f^{-1}(f(x)) = x,$$

$$\forall y \in Y, \quad (f \circ f^{-1})(y) = f(f^{-1}(y)) = y.$$

Hence $f^{-1} \circ f = \mathrm{Id}_X$ and $f \circ f^{-1} = \mathrm{Id}_Y$.
Assume that g is a correspondence such that $g \circ f = \mathrm{Id}_X$ and $f \circ g = \mathrm{Id}_Y$. Since identity correspondences are surjective mappings, by Proposition 3.5.4, we

deduce from the equality $g \circ f = \mathrm{Id}_X$ that g is surjective and $\mathrm{Dom}(f) = X = \mathrm{Im}(g)$. Similarly, we deduce from the equality $f \circ g = \mathrm{Id}_Y$ that f is surjective and $\mathrm{Dom}(g) = Y = \mathrm{Im}(f)$.
Since identity correspondences are injective, by Proposition 3.6.5, we deduce from $g \circ f = \mathrm{Id}_X$ that f is injective. Similarly, we deduce from $f \circ g = \mathrm{Id}_Y$ that f is a function. Therefore, f is a mapping which is injective and surjective, namely a bijection.
Finally, by Propositions 3.4.4 and 3.4.3, we obtain

$$g = g \circ \mathrm{Id}_Y = g \circ (f \circ f^{-1}) = (g \circ f) \circ f^{-1} = \mathrm{Id}_X \circ f^{-1} = f^{-1}.$$

□

Proposition 3.8.3 Let $f : X \to Y$ and $g : Y \to Z$ be bijections. Then the composite correspondence $g \circ f$ is also a bijection.

Proof This is a direct consequence of Propositions 3.7.5, 3.6.4 and 3.5.5

□

Proposition 3.8.4 Let X and Y be sets, f be a correspondence from X to Y, and g be a correspondence from Y to X. If $f \circ g$ and $g \circ f$ are bijections, then f and g are both bijections.

Proof By Proposition 3.5.4, f and g are surjective and are multivalued mappings. In particular,

$$\mathrm{Dom}(f) = X, \quad \mathrm{Im}(f) = Y, \quad \mathrm{Dom}(g) = Y, \quad \mathrm{Im}(g) = X.$$

Therefore, by Proposition 3.6.5, we deduce that f and g are injective and are functions. Hence f and g are both bijections.

□

3.9 Direct product

Definition 3.9.1 Let I be a set and $(A_i)_{i \in I}$ be a family of sets parametrised by I. We denote by

$$\prod_{i \in I} A_i$$

the set of all mappings from I to $\bigcup_{i\in I} A_i$ which send any $i \in I$ to an element of A_i. This set is called the **direct product** of $(A_i)_{i\in I}$. Using Notation 3.7.2 we often write an element of the direct product in the form of a family $x := (x_i)_{i\in I}$ parametrised by I, where each x_i is an element of A_i, called the i-th *coordinate* of x. In the case where I is the empty set, the union $\bigcup_{i\in I} A_i$ is empty. Therefore, the direct product contains a unique element (identity mapping of $\varnothing$).
For each $j \in I$, we denote by

$$\mathrm{pr}_j : \prod_{i\in I} A_i \longrightarrow A_j$$

the mapping which sends each element $(a_i)_{i\in I}$ of the direct product to its j-th coordinate a_j. This mapping is called the *projection to the j-th coordinate*.

Notation 3.9.2 Let n be a non-zero natural number. If $(A_i)_{i\in\{1,\ldots,n\}}$ is a family of sets parametrised by $\{1, \ldots, n\}$, then the set

$$\prod_{i\in\{1,\ldots,n\}} A_i$$

is often denoted as

$$A_1 \times \cdots \times A_n.$$

Axiom 1 (Axiom of choice) In this book, we adopt the following axiom. If I is a non-empty set and if $(A_i)_{i\in I}$ is a family of non-empty sets, then the direct product $\prod_{i\in I} A_i$ is not empty.

Proposition 3.9.3 Let I be a set and $(A_i)_{i\in I}$ be a family of sets parametrised by I. For any set X, the mapping

$$\Big(\prod_{i\in I} A_i\Big)^X \longrightarrow \prod_{i\in I} A_i^X,$$

which sends f to $(\mathrm{pr}_i \circ f)_{i\in I}$, is a bijection.

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "f_j"'] & \prod_{i\in I} A_i \arrow[d, "\mathrm{pr}_j"] \\ & A_j \end{tikzcd}$$

Proof Let $(f_i)_{i\in I}$ be an element of

$$\prod_{i\in I} A_i^X,$$

where each f_i is a mapping from X to A_i. Let $f : X \to \prod_{i\in I} A_i$ be the mapping which sends $x \in X$ to $(f_i(x))_{i\in I}$. By definition, for any $i \in I$ one has

$$\forall x \in X, \quad \mathrm{pr}_i(f(x)) = f_i(x).$$

Therefore the mapping is surjective.
If f and g are two mappings from X to $\prod_{i\in I} A_i$ such that $\mathrm{pr}_i \circ f = \mathrm{pr}_i \circ g$ for any $i \in I$, then, for any $x \in X$ one has

$$\forall i \in I, \quad \mathrm{pr}_i(f(x)) = \mathrm{pr}_i(g(x)).$$

Hence $f(x) = g(x)$ for any $x \in X$, namely $f = g$. Therefore the mapping is injective. □

Notation 3.9.4 Let I be a set, $(A_i)_{i\in I}$ be a family of sets parametrised by I. Let X be a set. For any $i \in I$, let $f_i : X \to A_i$ be a mapping from X to A_i. By Proposition 3.9.3 there exists a unique mapping $f : X \to \prod_{i\in I} A_i$ such that $\mathrm{pr}_i \circ f = f_i$ for any $i \in I$. By abuse of notation, we denote by $(f_i)_{i\in I}$ this mapping.
Let $(B_i)_{i\in I}$ be a family of sets parametrised by I. For any $i \in I$, let $g_i : B_i \to A_i$ be a mapping from B_i to A_i. We denote by

$$\prod_{i\in I} g_i : \prod_{i\in I} B_i \longrightarrow \prod_{i\in I} A_i$$

the mapping which sends $(b_i)_{i\in I}$ to $(g_i(b_i))_{i\in I}$. In the case where $I = \{1, \ldots, n\}$, where n is a non-zero natural number, the mapping $\prod_{i\in\{1,\ldots,n\}} g_i$ is also denoted as

$$g_1 \times \cdots \times g_n.$$

Proposition 3.9.5 Let $f : X \to Y$ be a mapping.

(1) If f is surjective, then there exists an injective mapping $g : Y \to X$ such that $f \circ g = \mathrm{Id}_Y$.

(2) If f is injective and X is not empty, then there exists a surjective mapping $h : Y \to X$ such that $h \circ f = \mathrm{Id}_X$.

Proof (1) The case where $Y = \varnothing$ is trivial since in this case $X = \varnothing$ and f is the identity mapping of $\varnothing$. In the following, we assume that Y is not empty. Since f is surjective, for any $y \in Y$, the set $f^{-1}(\{y\})$ is not empty. Hence the direct product

$$\prod_{y \in Y} f^{-1}(\{y\})$$

is not empty. In other words, there exists a mapping g from Y to X such that $f(g(y)) = y$ for any $y \in Y$, that is $f \circ g = \mathrm{Id}_Y$. By (2) of Remark 3.7.6 g is injective.
(2) Let x_0 be an element of X. We define a mapping $h : Y \to X$ as follows:

$$h(y) := \begin{cases} f^{-1}(y), & \text{if } y \in \mathrm{Im}(f), \\ x_0, & \text{else}. \end{cases}$$

Then, by construction one has $h \circ f = \mathrm{Id}_X$.
By (1) of Remark 3.7.6 h is surjective. □

3.10 Restriction and Extension

Definition 3.10.1 Let f and g be correspondence. If $\Gamma_f \subseteq \Gamma_g$, we say that f is a **restriction** of g and that g is an **extension** of f
Let X anf Y be sets, h be a correspondence from X to Y , and A be a subset of X.Denote by $h|_A$ the correspondence from A to Y such that

$$\Gamma_{h|_A} = \Gamma_h \bigcap (A \times Y).$$

We call it the **restriction of h to A**

Chapter 4

Binary Relations

†This chapter was first written in pre-course, then added some sections in make-up session, which titled "Ordering".Some sections have the same knowledge.It's a bit mess.

4.1 Generalities

Definition 4.1.1 Let X be a set , we call **binary relation** on X any correspondence from X to X .If R is a binary relation on X , for any $(x, y) \in X \times X$ we denote by xRy the statement $(x, y) \in \Gamma_R$.

Example 4.1.2 We denote by" $=$" the correspondence Id_X.

Definition 4.1.3 If R is a binary relation on X, we denote by $\not R$ the binary relation such that

$$x \not R y \Leftrightarrow (x, y) \notin \Gamma_R.$$

4.2 Equivalent Relation

Section 5.5:Quotient, will use this concept.

Definition 4.2.1 Let X be a set and R a binary relation on X.
(1) If $\forall x \in X, xRx$, we say that R is **reflexive**.
(2) If $\forall (x, y) \in X \times X, xRy \Rightarrow yRx$, we say that R is **symmetric**.
(3) If for all x, y, z of X, $xRy \wedge yRz \Rightarrow xRz$, we say that R is **transitive**.
(4) If R is reflexive, symmetric and transitive, we say that R is an **equivalent relation**.

Definition 4.2.2 Let $\sim$ be an equivalent relation on X. For any $x \in X$, we call the set

$$[x] := \{y \in X \mid y \sim x\}$$

the equivalent class of x under $\sim$, we denote by $X/\sim$ the set $\{[x] \mid x \in X\}$ of all equivalent class.It is a subset of $\mathscr{P}(X)$.Moreover, since $\forall x \in X, x \in [x]$, one has

$$X = \bigcup_{A \in X/\sim} A.$$

Proposition 4.2.3 $\forall (x, y) \in X \times X$, either $[x] = [y]$ or $[x] \cap [y] = \varnothing$.

Definition 4.2.4 The mapping $\pi : X \to X/\sim$ is called the **projection mapping** of $\sim$.

Proposition 4.2.5 (Theorem 5.5.5) $f : X \to Y$ be a mapping, if $\forall (x, y) \in X \times X, x \sim y \Rightarrow f(x) = f(y)$, then there exists a unique mapping

$$\tilde{f} : X/\sim \to Y, [x] \mapsto f(x),$$

such that

$$\tilde{f} \circ \pi = f.$$

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "\pi"'] & Y \\ X/\sim \arrow[ur, "\tilde{f}"'] & \end{tikzcd}$$

4.3 Partial Order

Definition 4.3.1 If
(1) R is reflexive.
(2) R is antisymmetric $\forall (x, y) \in X^2, xRy$ and yRx then $x = y$.
(3) R is transitive.
then we say that R is a **partial order** on X and (X, R) is a **partially ordered set**.If in addition , $\forall (x, y) \in X, xRy$ or yRx, we say that R is a **total order** and (X, R) is totally ordered set.

Example 4.3.2 $(\mathbb{R}, \leq)$ is a totally ordered set.$(\mathbb{N}, |)$ is a partially ordered set.

Definition 4.3.3 Let $(X, \underline{R})$ be a partially ordered set. We denote by R the binary relation on X defined as:

$$xRy \Leftrightarrow x\underline{R}y \wedge x \neq y,$$

we call R the **strict partial order**(not a partial order) associated with $\underline{R}$.

Example 4.3.4
(1) $<$ on $\mathbb{R}$.
(2) $\subset$ on $\mathscr{P}(X)$.

Proposition 4.3.5 R is the strict partial order associated with some partial order if and only if the following condition are satisfied:
(1) Irreflexivity $\forall x \in X, x\not{R}x$.
(2) Asymmetry.$\forall (x, y) \in X^2, xRy \Rightarrow y\not{R}x$.
(3) Transitivity.

Proof "$\Rightarrow$": easy.
"$\Leftarrow$":Suppose that R is a binary relation satisfying $(1) \sim (3)$. Define another binary relation $\underline{R}$ on X as:

$$x\underline{R}y \Leftrightarrow xRy \vee x = y.$$

We claim that $xRy \Leftrightarrow x\underline{R}y \wedge x \neq y$:
Suppose that xRy, then by definition, $x\underline{R}y$. By the irreflexivity, $x \neq y$.
Conversely, if $x\underline{R}y \wedge x \neq y$, then xRy should be true. $\square$

4.4 Monotonic Functions

Definition 4.4.1 Let $(I, \leq)$ and $(X, \leq)$ be partially ordered sets, and f be a function from I to X.
(1) If $\forall (x, y) \in \mathrm{Dom}(f)^2, x < y \Rightarrow f(x) \leq f(y)$ we say that f is increasing.
(2) If $\forall (x, y) \in \mathrm{Dom}(f)^2, x < y \Rightarrow f(x) < f(y)$, we say that f is strictly increasing.
(3) If $\forall (x, y) \in \mathrm{Dom}(f)^2, x < y \Rightarrow f(x) \geq f(y)$, we say that f is decreasing.

(4) If $\forall (x, y) \in \mathrm{Dom}(f)^2, x < y \Rightarrow f(x) > f(y)$, we say that f is strictly decreasing.
Increasing and decreasing functions are called **monotonic function**, strictly increasing and decreasing functions are called **strictly monotonic function**.

Proposition 4.4.2 Let f, g be functions between partially ordered sets.
(1) If both f and g are increasing or both f and g are decreasing, then $g \circ f$ is increasing.
(2) If one function between f and g is increasing while the order is decreasing, then $g \circ f$ is decreasing.

Proposition 4.4.3 Let f be a function between partially ordered set. If f is monotonic and injective, then f is strictly monotonic.

Proposition 4.4.4 Let I be a totally ordered set, X be a partially ordered set, and f be a function from I to X. If f is strictly monotonic, then f is injective.

Proof Let $(x, y) \in \mathrm{Dom}(f)^2$, such that $f(x) = f(y)$. Since I is totally ordered, then $x < y$ or $x > y$ or $x = y$. Suppose that f is strictly increasing. If $x < y$, then $f(x) < f(y)$, contradiction. If $x > y$, then $f(x) > f(y)$, contradiction. $\square$

Proposition 4.4.5 Let X be a totally ordered set, Y be an partially ordered set, f be an injective function from X to Y. If f is monotonic, then f^{-1} is also monotonic, and they have the same monotonic direction.

Proof We may suppose that f is increasing. Let $(a, b) \in \mathrm{Dom}(f^{-1})^2 = \mathrm{Im}(f)^2, a < b$. Since f^{-1} is a injective function, $f^{-1}(a) \neq f^{-1}(b)$, so either $f^{-1}(a) < f^{-1}(b)$ or $f^{-1}(a) > f^{-1}(b)$. If

$$f^{-1}(a) > f^{-1}(b), a = f(f^{-1}(a)) > f^{-1}(b) = b,$$

contradiction. Therefore, $f^{-1}(a) < f^{-1}(b)$. Hence f^{-1} is strictly increasing. $\square$

4.5 Bounds

Definition 4.5.1 Let $(X, \leq)$ be a partially ordered set , let A be a subset of X.
(1) Let $M \in X$. If $\forall a \in A, a \leq M$, we say that M is an upper bound of A.
(2) Let $m \in X$. If $\forall a \in A, m \leq a$, we say that m is an lower bound of A.
Denote by A^{u} the set of upper bounds of A in $(X, \leq)$.
Denote by A^{l} the set of lower bounds of A in $(X, \leq)$.

Example 4.5.2 $\Omega = \{1, 2, 3\}, X = \mathscr{P}(\Omega).(X, \subseteq)$ forms a partially ordered set. Let $A = \{\{1\}, \{2\}, \{1, 2\}\}, A^{\mathrm{u}} = \{\{1, 2\}, \{1, 2, 3\}\}, A^{\mathrm{l}} = \{\varnothing\}$.

Definition 4.5.3 Let $(X, \leq)$ be a partially ordered set , let A be a subset of X.
(1) If $M \in A$ is an upper bound of A, we say that M is the **greatest element** of A, denote as $\max_{\leq} A$.
(2) If $m \in A$ is an lower bound of A, we say that m is the **least element** of A, denote as $\min_{\leq} A$.
If there is not ambiguity on $\leq$, we can also write as $\max A, \min A$.

Definition 4.5.4 $A \subseteq Y \subseteq X$, let $A^{\mathrm{u}}_Y := \{y \in Y \mid \forall a \in A, a \leq y\}$ be the set of upper bounds of A in Y.If A^{u}_Y has a least element, we call it the **supremum** of A in Y, denoted as $\sup_{(Y,\leq)} A$, if there's no ambiguity on $\leq$ we can also write as $\sup_Y A$. Resp. **infimum**.

Notation 4.5.5 Let $(X, \leq)$ be a partially ordered set, $f : I \to X$ be a function.

$$\max f(I), \min f(I), \sup f(I), \inf f(I)$$

are written as

$$\max f, \min f, \sup f, \inf f.$$

Let $(X, \leq)$ be a partially ordered set, and $(x_i)_{i \in I} \in X^I$,

$$\max\{x_i | i \in I\}, \min\{x_i | i \in I\}, \sup\{x_i | i \in I\}, \inf\{x_i | i \in I\}$$

are denoted as

$$\max_{i \in I} x_i, \min_{i \in I} x_i, \sup_{i \in I} x_i, \inf_{i \in I} x_i.$$

Proposition 4.5.6
Let $(X, \leq)$ be a partially ordered set $(A, Z, Y) \in \mathscr{P}(X)^3, A \subseteq Z \subseteq Y$.
(1) If $\max A$ exists, then it is also the supremum of A in $(Y, \leq)$. Resp. infimum
(2) If $\sup_{(Y,\leq)} A$ exists and belongs to Z, then it is also the supremum of A in $(Z, \leq)$. Resp. infimum.

Proof
(1) By definition, $\max A$ is an upper bound of A. Since $A \subseteq Y, \max A \in Y$, Hence $\max A \in A^{\mathrm{u}}_Y$. Let $M \in A^{\mathrm{u}}_Y$. Since M is upper bound of A and $\max A \in A, \max A \leq M$. Then $\max A = \min A^{\mathrm{u}}_Y$.
(2) Since $Z \subseteq Y, A^{\mathrm{u}}_Z \subseteq A^{\mathrm{u}}_Y$. For any $M \in A^{\mathrm{u}}_Z$, one has $\sup_{Y,\leq} A \leq M$. If $\sup_{(Y,\leq)} A \in Z$, then $\sup_{(Y,\leq)} A \in A^{\mathrm{u}}_Z$. Hence $\sup_{(Y,\leq)} A = \min A^{\mathrm{u}}_Z$. □

Proposition 4.5.7
Let $(X, \leq)$ be a partially ordered set, $(A, B, Y) \in \mathscr{P}(X)^3, A \subseteq B \subseteq Y$.
(1) If $\sup_{(Y,\leq)} A$ and $\sup_{(Y,\leq)} B$ exist, then

$$\sup_{(Y,\leq)} A \leq \sup_{(Y,\leq)} B.$$

(2) If $\inf_{(Y,\leq)} A$ and $\inf_{(Y,\leq)} B$ exist, then

$$\inf_{(Y,\leq)} B \leq \inf_{(Y,\leq)} A.$$

Proof
(1) $\forall x \in A$, since $A \subseteq B, x \in B \leq \sup B$, by definition, $\sup B$ is an upper bound of A, $\sup B \in A^{\mathrm{u}}_Y$. $\sup A$ is the least in A^{u}_Y. Hence, $\sup_{(Y,\leq)} A \leq \sup_{(Y,\leq)} B$. □

Proposition 4.5.8 Let $(X, \leq)$ be a partially ordered set, f, g be elements of X^I where I is a set. Suppose that, $\forall i \in I, f(i) \leq g(i)$
(1) If $\sup f, \sup g$ exist, then $\sup f \leq \sup g$.
(2) Resp. infimum.

Proof $\forall t \in I, f(t) \leq g(t) \leq \sup g$, hence $\sup g$ is an upper bound of f. Since $\sup f$ is a the least upper bound of $f(i)$, $\sup f \leq \sup g$. □

Proposition 4.5.9 Let I be a totally ordered set $J \subseteq I$, and $f : I \to X$ be a mapping. Assume that J does not have any upper bound in I.
(1) If f is increasing, then $f(I)^{\mathrm{u}} = f(J)^{\mathrm{u}}$.
(2) If f is decreasing, then $f(I)^{\mathrm{l}} = f(J)^{\mathrm{l}}$.

Proof
(1) $f(J) \subseteq f(I)$ Any upper bound of $f(I)$ is also an upper bound of $f(J)$, hence $f(I)^{\mathrm{u}} \subseteq f(J)^{\mathrm{u}}$. Let $M \in f(J)^{\mathrm{u}}$, for any $i \in I, \exists j \in J, i < j$. Hence $f(i) \leq f(j) \leq M$. So $M \in f(I)^{\mathrm{u}}$, $f(J)^{\mathrm{u}} \subseteq f(I)^{\mathrm{u}}$. Therefore, $f(I)^{\mathrm{u}} = f(J)^{\mathrm{u}}$. $\square$

Proposition 4.5.10 Let $(X, \leq)$ be a partially ordered set , $Y \subseteq X, I$ be a set, and $(A_i)_{i\in I} \in \mathscr{P}(Y)^I$. Let $A = \bigcup\limits_{i\in I} A_i$
(1) Suppose that , $\forall i \in I, A_i$ has a supremum y_i in $(Y, \leq)$ and $\{y_i \mid i \in I\}$ has a supremum in $(Y, \leq)$. Then A has a supremum in $(Y, \leq)$ and

$$\sup_{(Y,\leq)} A = \sup_{(Y,\leq)} \{y_i \mid i \in I\}.$$

(2) Resp. inf.

Proof Let $y = \sup_{(Y,\leq)}\{y_i \mid i \in I\}, \forall a \in A, \exists i \in I, a \in A_i$. Hence $a \leq y_i \leq y$. Thus y is an upper bound of A in Y. Let $M \in A_Y^{\mathrm{u}}, \forall i \in I, M \in (A_i)_Y^{\mathrm{u}}$. So $y_i \leq M$. We then deduce that $y \leq M$. $\square$

Proposition 4.5.11 Let $(X, \leq)$ be a partially ordered set , $Y \subseteq X$.

$$\varnothing_Y^{\mathrm{u}} = \varnothing_Y^{\mathrm{l}} = Y.$$

4.6 Intervals

Definition 4.6.1 (Intervals) Let $(X, \leq)$ be a partially ordered set.$\forall (a, b) \in X^2$, let

$$[a, b] := \{x \in X \mid a \leq x \leq b\},$$

$$[a, b[:= \{x \in X \mid a \leq x < b\}.$$

We say that a subset is a **interval** if $\forall (a, b) \in I^2, [a, b] \subseteq I$.

Proposition 4.6.2 Let $(X, \leq)$ be a partially ordered set, let Λ be a non-empty set and $(I_\lambda)_{\lambda\in\Lambda}$ be a family of interval in X, then
(1) $I := \bigcap\limits_{\lambda\in\Lambda} I_\lambda$ is an intervals.
(2) If $\bigcap\limits_{\lambda\in\Lambda} I_\lambda \neq \varnothing$, then $J := \bigcup\limits_{\lambda\in\Lambda} I_\lambda$ is an interval.

Proof
(2): Let $x \in I = \bigcap_{\lambda\in\Lambda} I_\lambda$, let $(a, b) \in J^2, \exists(\alpha, \beta) \in \Lambda^2, \alpha \in I_\alpha, \beta \in I_\beta$.We will show that $[a, b] \subseteq I_\alpha \cup I_\beta$. If $a \not\leq b$, then $[a, b] \neq \varnothing \subseteq I_\alpha \cup I_\beta$. We may assume $a \leq b$.
If $b \leq x$, then $[a, b] \subseteq [a, x] \subseteq I_\alpha$, if $x \leq a$, then $[a, b] \subseteq [x, b] \subseteq I_\beta$. Suppose that $a < x < b$, one has $[a, b] = [a, x] \cup [x, b]$ and so on , $[a, b] = [a, x] \cup [x, b] \subseteq I_\alpha \cup I_\beta \subseteq J$.

□

Definition 4.6.3 (Endpoints)
Let $(X, \leq)$ be a partially ordered set and I be a non-empty interval in X.
If $\sup I$ exists, we call it the right endpoint of I.
If $\inf I$ exists, we call it the left endpoint of I.

Proposition 4.6.4 Let $(X, \leq)$ be a totally ordered set and I be a interval in X
(1) Suppose that I has a supremum b in $X, \forall x \in I, [x, b[\subseteq I$.
(2) Suppose that I has a infimum b in $X, \forall x \in I,]b, x] \subseteq I$.

Proposition 4.6.5 Let $(X, \leq)$ be a totally ordered set and I be a non-empty interval in X. Assume that I has an infimum a and a supremum b in X. Then I is one of the following sets:$[a, b],\ [a, b[,\ ]a, b],\ ]a, b[$.

Proof $\forall x \in I, a \leq x \leq b$, hence $I \subseteq [a, b]$.
(i) if $\{a, b\} \in I$, then $I = [a, b]$.
(ii) if $a \in I, b \notin I, I \subseteq [a, b[= [a, b]\backslash\{b\}$. Let $x \in [a, b[$, since $x < b$, x is not an upper bound of I. Hence $\exists y \in I, x < y$. Note that $[a, y] \subseteq I$, hence $x \in I$, therefore $[a, b[\subseteq I$. Similarly , is $b \in I, a \notin I$, then $]a, b] = I$.
(iii) if $\{a, b\} \cap I = \varnothing$, then $I \subseteq]a, b[$. $\forall x \in]a, b[, \exists s, t \in I, s < x < t$. Hence $x \in [s, t] \subseteq I$. Therefore $]a, b[= I$.

□

Definition 4.6.6 (Dense) Let $(X, \leq)$ be a totally ordered set, if $\forall (x, z) \in X^2, x < z \Rightarrow]x, z[\neq \varnothing$ then we say that $(X, \leq)$ is **dense**.

Proposition 4.6.7 Let $(X, \leq)$ be a totally ordered set that is dense, $(a, b) \in X^2, a < b$. If I is one of the intervals $[a, b], [a, b[\ldots$, then $a = \inf I, b = \sup I$.

Proof By definition, b is an upper bound of I, since $(X, \leq)$ is a totally ordered set, if b is not the supremum of I , $\exists M \in I^{\mathrm{u}}$ such that $M < b$. Let $x \in I$, one has $x \leq M < b$. Since $[x, b[\subseteq I, M \in I$, hence $M = \max I$. Since X is dense , pick $M' \in]M, b[$. Since $M \in I, b = \sup I, [M, b[\subseteq I$. Hence $M' \in I, M' \leq M$. This contradicts $M < M'$. □

4.7 Well-ordered Set

Definition 4.7.1 (Well-ordered) Let $(X, \leq)$ be a partially ordered set. If $\forall A \in \mathscr{P}(X), A \neq \varnothing \Rightarrow A$ has a least element, we say that $(X, \leq)$ is a **well-ordered set**.

Axiom 2 $(\mathbb{N}, \leq)$ is a well-ordered set.

Proposition 4.7.2 If $(X, \leq)$ is a well-ordered set, then it is a totally ordered set.

Proposition 4.7.3 $(X, \leq)$ is a well-ordered set, $Y \subseteq X$, then $(Y, \leq)$ is a well-ordered set.

Theorem 4.7.4 Let $(X, \leq)$ be a well-ordered set. Let $P(\cdot)$ be a condition on X. If

$$\forall x \in X, (\forall y \in X_{<x}, P(y)) \Rightarrow P(x),$$

then $\forall x \in X, P(x)$.

Remark 4.7.5 Suppose that $X \neq \varnothing$, there is a least element m of X. The statement

$$(\forall y \in X_{<m}, P(y)) \Rightarrow P(m) \text{ and } P(m) \text{ have the same truth value.}$$

Proof Let $A = \{x \in X \mid \neg P(x)\}$. If $A \neq \varnothing, \exists x \in A$ which is the least element of A. By definition, $(\forall y \in X_{<x}, P(y))$ is true. It contradicts to . $\square$

Remark 4.7.6 We add a formal element $+\infty$ to $\mathbb{N}$ and require $\forall n \in \mathbb{N}, n < +\infty$.

Fact: $\mathbb{N} \cup \{+\infty\}$ is a well-ordered set. Let $P(\cdot)$ be a condition on $\mathbb{N} \cup \{+\infty\}$. We need to check:

1. $P(0)$.
2. $\forall n \in \mathbb{N}_{\leq 1}, P(0) \wedge \cdots \wedge P(n-1) \Rightarrow P(n)$.
3. $(\forall n \in \mathbb{N}, P(n)) \Rightarrow P(+\infty)$.

4.8 Order-completeness

Definition 4.8.1 Let $(X, \leq)$ be a partially ordered set. If any subset of X has a supremum in X, we say that $(X, \leq)$ is **order-complete**. Note that an order-complete partially ordered set is never empty.

Axiom 3 Let $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$, where $-\infty, +\infty$ are distinct formal elements that do not belongs to $\mathbb{R}$. If we equip $\overline{\mathbb{R}}$ with the total order extending that of $\mathbb{R}$ such that

$$\forall x \in \mathbb{R}, -\infty < x < +\infty,$$

then $(\overline{\mathbb{R}}, \leq)$ is order complete.

Example 4.8.2 Let Ω be a set, $X = \mathscr{P}(\Omega)$. Then $(X, \subseteq)$ is order complete.

Proof Let $Y \subseteq X$. Then

$$Y^{\mathrm{u}} = \{B \in \mathscr{P}(\Omega) \mid \forall A \in Y, A \subseteq B\}.$$

$\bigcup\limits_{A \in Y} A$ is the least upper bound of Y in X. So $\sup(Y) = \bigcup\limits_{A \in Y} A$. $\square$

Proposition 4.8.3 Let $(X, \leq)$ be an order complete partially ordered set. Any subset of X has an infimum in X.

Proof Let $A \subseteq X,\ m := \sup A^{\mathrm{l}}$. We prove that $m \in A^{\mathrm{l}}$.
Let $x \in A, \forall y \in A^{\mathrm{l}},\ y \leq x$, so $x \in (A^{\mathrm{l}})^{\mathrm{u}}$. Hence $m \leq x$. □

Here Huayi gave a notation which have been given in Notation4.5.5, then came to Proposition4.5.6 and the following.

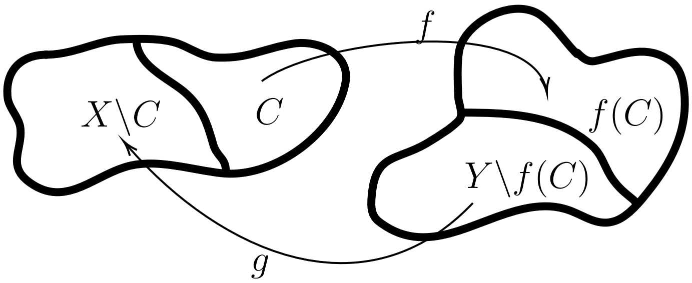
Definition 4.8.4 Let X be a set and $f : X \to X$ be a mapping. If $x \in X$ is such that $f(x) = x$, then we say that x is a fixed point of f.

Theorem 4.8.5 (Knaster-Tarski fixed point)
Let $(X, \leq)$ be an order complete partially ordered set , $f : X \to X$ be an increasing mapping. Let
$$F = \{x \in X \mid f(x) = x\},$$
then $(F, \leq)$ is order complete. In particular $F \neq \varnothing$.

Proof Let A be a subset of F. We consider
$$S_A := \{y \in A^{\mathrm{u}} \mid f(y) \leq y\}.$$
Let $m := \inf S_A, \forall a \in A, a$ is a lower bound of S_A. So $a \leq m$. So $m \in A^{\mathrm{u}}, \sup A \leq m$. For any $y \in S_A$, one has $m \leq y$. Since f is increasing, $f(m) \leq f(y) \leq y$. So $f(m)$ is a lower bound of S_A, which leads to $f(m) \leq m$. That means $m \in S_A$. Hence $m = \min S_A$. For any $x \in A, x = f(x) \leq f(m)$. So $f(m) \in A^{\mathrm{u}}$. Moreover, since $f(m) \leq m, f(f(m)) \leq f(m)$. So $f(m)$ is an element of S_A, which leads to $m \leq f(m)$. Hence $m \in F$. Therefore, $m = \sup_{(F,\leq)} A$. □

Definition 4.8.6 Let X, Y be sets. If there exists a bijection from X to Y, we say that X and Y are **equipotent**.

Theorem 4.8.7 (Cantor-Bernstein) Let X and Y be sets. Assume that there exists injective mappings $f : X \to Y$ and $g : Y \to X$. Then X and Y are equipotent.



Proof Consider $\Phi : \mathscr{P}(X) \to \mathscr{P}(X), A \mapsto X\backslash g(Y\backslash f(A))$. If $(A, B) \in \mathscr{P}(X)^2$ such that $A \subseteq B$, then $f(A) \subseteq f(B), Y\backslash f(A) \supseteq Y\backslash f(B), g(Y\backslash f(A)) \supseteq g(Y\backslash f(A)), \Phi(A) \subseteq \Phi(B)$. So Φ is increasing. By Knaster-Tarski theorem, $\exists C \in \mathscr{P}(X), C = \Phi(C)$. Then $h : X \to Y, h(x) := \begin{cases} f(x), x \in C \\ g^{-1}(x), x \in X\backslash C \end{cases}$ is a bijection. □

Lemma 4.8.8 Let $(X, \leq)$ is a partially ordered set.
(1) Let $(A, B) \in \mathscr{P}(X)^2$, if $A \subseteq B$, then $B^{\mathrm{u}} \subseteq A^{\mathrm{u}}, B^{\mathrm{l}} \subseteq A^{\mathrm{l}}$.
(2) $\forall A \in \mathscr{P}(X), A \subseteq (A^{\mathrm{u}})^{\mathrm{l}} \cap (A^{\mathrm{l}})^{\mathrm{u}}$.

Theorem 4.8.9 (Dedekind-MacNeille)
Let $(X, \leq)$ be a partially ordered set. Let $\hat{X} := \{A \in \mathscr{P}(X) \mid (A^{\mathrm{u}})^{\mathrm{l}} = A\}$.
(1) $(\hat{X}, \subseteq)$ is order complete.
(2) $\forall A \in \mathscr{P}(X), A^{\mathrm{l}} \in \hat{X}$.
(3) $X \to \hat{X}, x \mapsto \{x\}^{\mathrm{l}}$ is strictly increasing.
(4) $\forall A \in \hat{X}$ one has $A = \cup_{x\in A}\{x\}^{\mathrm{l}} = \cup_{x\in A}\hat{x}$. In particular,

$$A = \sup_{(\hat{X},\subseteq)} \{\hat{x} \mid x \in A\}.$$

(5) Let $A \in \hat{X}$. If $A^{\mathrm{u}} = \varnothing$, then $A = X$. If $A^{\mathrm{u}} \neq \varnothing$, then

$$A = \bigcap_{x\in A^{\mathrm{u}}} \hat{x} = \inf_{(\hat{X},\subseteq)} \{\hat{x} \mid x \in A^{\mathrm{u}}\},$$

$$A = \bigcup_{x\in A} \hat{x} = \sup_{(\mathscr{P}(X),\subseteq)} \{\hat{x} \mid x \in A\} = \sup_{(\hat{X},\subseteq)} \{\hat{x} \mid x \in A\}.$$

Remark 4.8.10 We've know that $(\mathscr{P}(X), \subseteq)$ is order complete. So for the sets not order complete, we can build a relation between them to make it become order complete. And this theorem tell us how to do.

Proof
(1) Consider $\Phi : \mathscr{P}(X) \to \mathscr{P}(X), A \mapsto (A^{\mathrm{u}})^{\mathrm{l}}$. By the lemma, Φ is increasing. Since $\mathscr{P}(X)$ is complete, and $\hat{X}$ is the set of fixed point of Φ. By Knaster-Tarski fixed point theorem, $(\hat{X}, \subseteq)$ is order complete.
(2) Let $A \in \mathscr{P}(X)$, we prove that $A^{\mathrm{l}} = ((A^{\mathrm{l}})^{\mathrm{u}})^{\mathrm{l}}$. Since $A \subseteq (A^{\mathrm{l}})^{\mathrm{u}}$ (by the lemma), $((A^{\mathrm{l}})^{\mathrm{u}})^{\mathrm{l}} \subseteq A^{\mathrm{l}}$, by (2) of the lemma applied to A^{l}. Hence $A^{\mathrm{l}} = ((A^{\mathrm{l}})^{\mathrm{u}})^{\mathrm{l}}$
(3) Let x and y be element of X such that $x < y$ then $\{x\}^{\mathrm{l}} \subsetneq \{y\}^{\mathrm{l}}$. In fact, if $z \in \{x\}^{\mathrm{l}}, z \leq x$. Since $x < y, z < y$. Moreover, $y \in \{y\}^{\mathrm{l}}$, but $y \notin \{x\}^{\mathrm{l}}$.
(4) $\forall x \in A, x \in \{x\}^{\mathrm{l}} = \hat{x}$. So $A \subseteq \bigcup_{x \in A} \hat{x}$. Conversely, $\forall x \in A, x = \min(\{x\}^{\mathrm{u}})$. Hence $\{x\}^{\mathrm{l}} = (\{x\}^{\mathrm{u}})^{\mathrm{l}} \subseteq (A^{\mathrm{u}})^{\mathrm{l}} = A$. Therefore $\bigcup_{x \in A} \{x\}^{\mathrm{l}} \subseteq A$. Finally we get $\bigcup_{x \in A} \hat{x} = A \in \hat{X}$.
(5) If $A^{\mathrm{u}} = \varnothing$ then $A = (A^{\mathrm{u}})^{\mathrm{l}} = \varnothing^{\mathrm{l}} = X$. We assume that $A^{\mathrm{u}} \neq \varnothing$.

$$\inf_{(\mathscr{P}(X), \subseteq)} \{\hat{x} \mid x \in A^{\mathrm{u}}\} = \bigcap_{x \in A^{\mathrm{u}}} \hat{x} = \bigcap_{x \in A^{\mathrm{u}}} \{x\}^{\mathrm{l}} = (A^{\mathrm{u}})^{\mathrm{l}} = A.$$

So it is equal to $\inf_{(\hat{X}, \subseteq)} \{\hat{x} \mid x \in A^{\mathrm{u}}\}$. □

Remark 4.8.11 $\forall A \in \hat{X}, A = \{x \in X \mid \hat{x} \subseteq A\}, A^{\mathrm{u}} = \{x \in X \mid A \subseteq \hat{x}\}$.

Definition 4.8.12 $\hat{X}$ is called the Dedekind-MacNeille order completion of $(X, \leq)$.

4.9 Recursive Construction

Definition 4.9.1 Let $(X, \leq)$ be a partially ordered set. Let $I \subseteq X$. If $\forall a \in I, X_{<a} \subseteq I$, we say that I is an initial segment of X.

Proposition 4.9.2 Let $(X, \leq)$ be a totally ordered set, I, J be initial segments of X. Either $I \subseteq J$ or $J \subseteq I$.

Proof Assume that $I\backslash J \neq \varnothing$, take $x \in I\backslash J$, for any $y \in J$, if $y \not\leq x$, then $x < y$ and hence $x \in X_{<y} \subseteq J$, contradiction. Therefore $y \leq x$. Then $y = x \in I$ or $y \in X_{<x} \subseteq I$. $\square$

Proposition 4.9.3 Let $(X, \leq)$ be a well-ordered set. I be an initial segment of X, such that $I \neq X$. There is a unique $a \in X$ such that $I = X_{<a}$.

Proof $X\backslash I \neq \varnothing$ Let $a = \min(X\backslash I)$. By definition, $I \subseteq X_{<a}$. In fact, $\forall y \in I$ if $y \nless a$, then $a \leq y$. Since I is an initial segment $a \in I$, contradiction.
Conversely, if $x \in X_{<a}$, then $x \notin X\backslash I$. Since otherwise $a \leq x$. Therefore $x \in I$.
Uniqueness, $\forall a \in X, a = \min(X\backslash X_{<a}) = \min(X_{\leq} a)$. Hence $X_{<a} = X_{<b} \Rightarrow a = b$. $\square$

Proposition 4.9.4 Let $(X, \leq)$ be a partially ordered set, Λ be a non-empty set, and $(I_\lambda)_{\lambda\in\Lambda}$ be a family of initial segments of X. Then

$$I := \bigcap_{\lambda\in\Lambda} I_\lambda, J := \bigcup_{\lambda\in\Lambda} I_\lambda$$

are initial segments of X.

Proof
Let $a \in I$. $\forall\lambda \in \Lambda, a \in I_\lambda$ and hence $X_{<a} \subseteq I_\lambda$. Therefore, $X_{<a} \subseteq \bigcap_{\lambda\in\Lambda} I_\lambda = I$.
Let $b \in J$. Then $\exists\lambda_0 \in \Lambda$ such that $b \in I_{\lambda_0}$. So $X_{<b} \subseteq I_{\lambda_0} \subseteq \bigcup_{\lambda\in\Lambda} I_\lambda = J$. $\square$

Theorem 4.9.5 (Recursive construction) Let $(X, \leq)$ be a well ordered set, and Y be a set. For any $x \in X$ and any mapping $h : X_{<x} \to Y$, we fix an element $\Phi(h) \in Y$. Then, there exists a unique mapping $f : X \to Y$ such that

$$\forall x \in X, f(x) = \Phi(f\mid_{X_{<x}}).$$

Example 4.9.6 For any $(a_0, \dots, a_{n-1}) \in \mathbb{R}^n$, we fix an element $a_{n-1} + \varepsilon \in \mathbb{R}$, where ε is a real number. There exists a unique mapping $(n \in \mathbb{N}) \mapsto f(n)$ such that $f(n) = f(n-1) + \varepsilon$. ($f(n) := n\varepsilon$).

Proof
"Uniqueness":
Let f, g be mappings from X to Y such that

$$\forall x \in X, f(x) = \Phi(f\,|_{X_{<x}}), g(x) = \Phi(g\,|_{X_{<x}}).$$

Then: $\forall x \in X$, we have

$$(\forall y \in X_{<x}, f(y) = g(y)) \Rightarrow f(x) = g(x).$$

So by inclusion $\forall x \in X, f(x) = g(x)$, namely, $f = g$.
"Existence":
Let $\mathscr{S}$ be the set of initial segments S of X such that $\exists f_S : S \to Y$ satisfying

$$\forall x \in S, f_S(x) = \Phi(f_S\,|_{X_{<x}}). \tag{$*$}$$

Let $X_0 = \cup_{S\in\mathscr{S}} S$. It is also an initial segment of X. For any $x \in X_0$ there exists S such that $x \in S$. If S_1 and S_2 are two elements of $\mathscr{S}$, then $S_1 \cap S_2$ is also an initial segment. Moreover $f_{S_1}\,|_{S_1\cap S_2}$ and $f_{S_2}\,|_{S_1\cap S_2}$ satisfy $(*)$ So $f_{S_1}\,|_{S_1\cap S_2}= f_{S_2}\,|_{S_1\cap S_2}$. Thus $f_S(x)$ does not depend on the choice of $S \in \mathscr{S}$ containing x. We denote it as $f(x)$. $f : X_0 \to Y$ satisfying $(*)$. So $X_0 \in \mathscr{S}$. If $X_0 \neq X.\exists a \in X$ such that $X_0 = X_{<a}$. We extend f to $X_0 \cup \{a\}$ by letting $f(a) = \Phi(f)$. Then we get $X_0 \cup \{a\} \in \mathscr{S}$.Contradiction.Therefore $X_0 = X$ and we get the existence of f. □

Definition 4.9.7 Let A be a set. If there exists an injective mapping $A \to \mathbb{N}$, then we say that A is **countable**. If there exists an injective mapping $f : A \to \mathbb{N}$ such that $f(A)$ is bounded from above (having an upper bound in $\mathbb{N}$), then we say that A is **finite**.

Lemma 4.9.8
(1) Let $n \in \mathbb{N}$ and $x_0, \ldots, x_n$ be elements of $\mathbb{N}$ such that $x_0 < \cdots < x_n$, then $\forall i \in \mathbb{N}_{\le n}, i \le x_i$.
(2) Let $(x_n)_{n\in\mathbb{N}}$ be a family of elements in $\mathbb{N}$ such that $\forall n \in \mathbb{N}, x_n < x_{n+1}$, then $\forall i \in \mathbb{N}, i \le x_i$.

Proof If $j \le x_j$ for $j \in \{0, \ldots, i-1\}$. Then, in the case where $i = 0, 0 \le x_0$ holds since $0 = \min_{\le} \mathbb{N}$. In the case where $i > 0$, one has $i - 1 \le x_{i-1} < x_i$. So $x_i \ge x_{i-1} + 1 \ge i - 1 + 1 = i$. □

Proposition 4.9.9 Let $f : A \to B$ be a mapping.
(1) If f is injective and if B is finite, then A is finite. (resp. countable)
(2) If f is surjective and A is finite, then B is finite. (resp countable)

Proof
(1) Let $g : B \to \mathbb{N}$ injective and bounded from above. Then $g \circ f$ is injective and $\mathrm{Im}(g \circ f) \subseteq \mathrm{Im}(g)$.
(2) $\exists$ injective mapping $B \to A$ by the axiom of choice. $f : A \to B$ For any $b \in B$, pick $h(b) \in f^{-1}(\{b\}) \subseteq A$, $h : B \to A$. If $h(b) = h(b')$, then $f(h(b)) = f(h(b')) = b'$. □

Proposition 4.9.10 Let X, Y be sets.
(1) If X and Y are finite, then $X \cup Y$ is finite.(resp. countable)
(2) If X is infinite and Y is finite, then $X \backslash Y$ is infinite.(resp. uncountable)

Proof
(1) Let $f : X \to \mathbb{N}$ and $g : Y \to \mathbb{N}$ be injective mappings. We construct $h : X \cup Y \to \mathbb{N}$ such that

$$h(x) = \begin{cases} 2f(x) & x \in X \\ 2g(x) + 1 & x \in Y \backslash X \end{cases}$$

h is then injective, and h is bounded if f and g are bounded.
In fact, if $(x, y) \in (X \cup Y)^2$,
either $(x, y) \in X^2$ and $h(x) = 2f(x) = h(y) = 2f(y)$ if and only if $x = y$.
or $(x, y) \in (Y \backslash X)^2$ and $h(x) = h(y) \Rightarrow x = y$.
or $x \in X, y \in Y \backslash X. h(x) \neq h(y)$ (So $h(x) = h(y) \Rightarrow x = y$).
or $y \in X, x \in Y \backslash X, h(x) \neq h(y)$.
(2) Assume that $X \backslash Y$ is finite, then $X = (X \backslash Y) \cup Y$ is also finite. □

Notation 4.9.11 If $f : X \to X$ is a mapping. Then f^0 denotes Id_X. For $n \in \mathbb{N}_{\geq 1}$, f^n denotes $\underbrace{f \circ f \circ \cdots \circ f}_{n}$.

Theorem 4.9.12 $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N}$ are equipotent.

Proof Let $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, $(a, b) \mapsto 2^a(2b+1)$. It is an injective mapping since $2^a(2b+1) = 2^{a'}(2b'+1)$. So $a = a', b = b'$. Moreover $x \mapsto (0, x)$ is injective. □

Corollary 4.9.13 $\forall n \in \mathbb{N}, n \geq 1, \mathbb{N}^n$ and $\mathbb{N}$ are equipotent.

Proof Induction on n.
For $n = 1$, easy. We assume that $\mathbb{N}^n$ is equipotent to $\mathbb{N}$ and $f : \mathbb{N}^n \to \mathbb{N}$ be a bijection. Then the mapping

$$f' : \mathbb{N}^n \times \mathbb{N} \to \mathbb{N} \times \mathbb{N}, \ (x_1, \ldots, x_n; x_{n+1}) \mapsto (f(x_1, \ldots, x_n), x_{n+1})$$

is a bijection. By Theorem 4.9.12, there exists a bijection $g : \mathbb{N}^2 \to \mathbb{N}$.Therefore,

$$g \circ f' : \mathbb{N}^{n+1} \to \mathbb{N}$$

is a bijection, which leads to $\mathbb{N}^{n+1}$ and $\mathbb{N}$ are equipotent. □

Motivation: Let X be a set. A sequence in X is by definition a family $(x_i)_{i\in I}$, where I is an infinite subset of $\mathbb{N}$, and each x_i is an element of X.

Example 4.9.14 $(a + bn)_{n\in\mathbb{N}}$; $\left(\frac{1}{n}\right)_{n\in\mathbb{N}_{\geq 1}}$.

Proposition 4.9.15 Let $I \subseteq \mathbb{N}$.
(1) $\mathrm{Id}_I : I \to I$ is the only increasing mapping bijection from I to I.
(2) If I is bounded from above, then Id_I is the only strictly increasing mapping from I to I.

Proof
(1) Let $f : I \to I$ be an increasing bijection. We want to prove:

$$A := \{x \in I \mid f(x) \neq x\} = \varnothing.$$

If this set is non-empty, it has a least element n_0. By definition, $f(n_0) \neq n_0$. So either $n_0 < f(n_0)$ or $n_0 > f(n_0)$.
If $f(n_0) < n_0$, then $f(n_0) \notin A$, and hence $f(n_0) = f(f(n_0)) < f(n_0)$, contradiction. So $n_0 < f(n_0)$. For any $n \in I$, if $n_0 \leq n$ then $f(n_0) \leq f(n)$. If $n_0 > n$, then $n \notin A$ and $f(n) = n < n_0$.(∗). Hence $f(n) \neq n_0$ for any $n \in I$. This contradicts the assumption that f is bijective.
(2) Suppose that I is bounded from above, and $f : I \to I$ is strictly increasing. We follow the same reasoning until (∗). $n_0 < f(n_0)$ implies that $\forall k \in \mathbb{N}$, $f^k(n_0) < f^{k+1}(n_0)$, that means

$$n_0 < f(n_0) < \cdots < f^{k+1}(n_0).$$

So by the lemma 4.9.8, $k \leq f^k(n_0)$, this contradicts the assumption that I is bounded from above. $\square$

Corollary 4.9.16 Let $I \subseteq \mathbb{N}$ bounded from above, and $J \subseteq I$. If $J \neq I$, there does not exist a strictly increasing mapping from I to J.

Proof Suppose that $f : I \to J$ is a strictly increasing mapping. Let $g : J \to I, x \mapsto x$ be a inclusion mapping. So $g \circ f : I \to I$ is strictly increasing and hence $g \circ f = \mathrm{Id}_I$. However $\mathrm{Im}(g \circ f) \subseteq \mathrm{Im}(g) = J \neq I$. Contradiction. $\square$

Proposition 4.9.17 Let $I \subseteq \mathbb{N}$ non-empty.
(1) If I is bounded from above, then there exists a unique pair (N, f), where $N \in \mathbb{N}$ and $f : \{0, 1, \ldots, N\} \to I$ is an increasing bijection. (We say that the cardinality of I is $N + 1$.)
(2) If I is NOT bounded from above, there exists an increasing bijection from $\mathbb{N}$ to I. (We say that the cardinality of I is $\aleph_0$)

Proof We construct in a recursive way a family of elements in I. Let $x_0 = \min(I)$. If $x_0, \ldots, x_n$ are chosen (with $x_0 < \cdots < x_n$) we pick $x_{n+1} = \min(I \backslash \{x_0, \ldots, x_n\})$. We slop at N if $\{x_0, \ldots, x_N\} = I$. Thus we obtain the increasing bijection needed by the proposition.
"Uniqueness" for (2): If $f : \mathbb{N} \to I, g : \mathbb{N} \to I$ are increasing bijections, then $f^{-1} \circ g : \mathbb{N} \to \mathbb{N}$ and $g^{-1} \circ f : \mathbb{N} \to \mathbb{N}$ are increasing bijections. So $f^{-1} \circ g = \mathrm{Id}_{\mathbb{N}}$. Hence $f = g$.
"Uniqueness" for (1): Let $f : \{0, 1, \ldots, N\} \to I$ and $g : \{0, 1, \ldots, M\} \to I$ be increasing bijections. $g^{-1} \circ f : \{0, 1, \ldots, N\} \to \{0, 1, \ldots, M\}$ and $g : \{0, 1, \ldots, M\} \to I$ be increasing bijections. $f^{-1} \circ g : \{0, 1, \ldots, M\} \to \{0, 1, \ldots, N\}$ are increasing bijection. So $N \leq M$ and $M \leq N$, which leads to $N = M, g = f$. $\square$

Corollary 4.9.18 A non-empty set X is finite if and only if it can be written as $\{x_0, \ldots, x_N\}$ where $N \in \mathbb{N}$, and $x_0, \ldots, x_N$ are distinct elements of X.

Proof Let $f : X \to \mathbb{N}$ be an injective mapping with $f(x)$ bounded from above. Then there exists (N, g) where $N \in \mathbb{N}$ and $g : \{0, \ldots, N\} \to f(x)$ is an increasing bijection. Then $f^{-1} \circ g : \{0, \ldots, N\} \to X$ is a bijection. We take x_i to be $(f^{-1} \circ g)(i)$ (Note that N is unique $N + 1$ is called the cardinality of X). $\square$

Proposition 4.9.19 Let X be a set. The following condition are equipotent:
(1) X is infinite.
(2) $\exists \mathbb{N} \to X$ injective.
(3) $\exists$ injective mapping $f : X \to X$ such that $f(X) \neq X$.

Proof
(1)⇒(2) We construct a sequence $(x_n)_{n\in\mathbb{N}}$ in X as follows. $X \neq \varnothing$. We pick arbitrary $x_0 \in X$. Suppose that distinct elements $x_0, \ldots, x_n$ of X are chosen. The set $X\backslash\{x_0, \ldots, x_n\} \neq \varnothing$ since otherwise $X = \{x_0, \ldots, x_n\}$ is finite. We pick $x_{n+1} \in X\backslash\{x_0, \ldots, x_n\}$, $x_0, \ldots, x_{n+1}$ are distinct. The mapping $\mathbb{N} \to X, n \mapsto x_n$ is injective.
(2)⇒(3) Let $f : \mathbb{N} \to X$ be injective. We define $g : X \to X$.

$$g(x) := \begin{cases} f(n+1) & , \quad x = f(n) \\ x & , \quad x \notin f(\mathbb{N}). \end{cases}$$

$g(X) \neq X$ since $f(0) \notin g(X)$ If $x \notin f(\mathbb{N}), g(x) = x \notin f(\mathbb{N})$, so $g(x) \neq f(0)$. If $x = f(n), g(x) = f(n+1) \neq f(0)$ since f is injective.
(3)⇒(2) Let $g : X \to X$ be injective with $g(X) \neq X$. We pick $x_0 \in X\backslash g(X)$. We define a sequence $(x_n)_{n\in\mathbb{N}}$ by letting $x_{n+1} := g(x_n)$. Since g is injective, $x_n \in g^n(X)\backslash g^{n+1}(X)$. Otherwise $\exists y \in X$, such that $x_n = g^n(x_0) = g^{n+1}(y)$. Hence $x_0 = g(y) \in g(x)$ contradiction. Then $x_0, x_1, \ldots,$ are distinct, which defines an injective mapping $\mathbb{N} \to X, n \mapsto x_n$.
(2)⇒(1) If X is finite, $\exists g : X \to \mathbb{N}$ injective with $g(x)$ bounded. Then $\mathbb{N} \to X \xrightarrow{g} \mathbb{N}$ is injective with $h(\mathbb{N})$ bounded from above. $\square$

Chapter 5

Groups

5.1 Composition Law

Definition 5.1.1 Let X be a set.

(i) A **compositon law** on X is a mapping

$$*: X \times X \to X, (x, y) \mapsto x * y.$$

(ii) Let $Y \subseteq X$ be a set , Y is **close under** $*$ if $\forall x, y \in Y, x * y \in Y$.

(iii) $*$ is **communitative** if $\forall (x, y) \in X^2, x * y = y * x$.

(iv) $*$ is **associative** if $\forall (x, y, z) \in X^3$, $(x*y)*z = x*(y*z)$. If $*$ is associative, then we can define

$$x_1 * x_2 * \cdots * x_n = (x_1 * x_2 * \cdots * x_{n-1}) * x_n.$$

(v) Let G be a set , $*$ is a composition law on G. If $*$ is associative, then we say $(G, *)$ is a **semigroup**.

Example 5.1.2
(1) Let $(X, *)$ be a composition law .We define $(X, \hat{*})$ satisfies:

$$\hat{*}: X \times X \to X, (x, y) \mapsto y * x.$$

By definition, $x = \hat{x} \Leftrightarrow *$ is communitative. If $*$ is associative, then so does $\hat{*}$. Let $\mathfrak{M}_X$ the set of all mapping from X to X. On $\mathfrak{M}_X$, the composition of mapping defines a composition law:

$$\begin{aligned}\mathfrak{M}_X \times \mathfrak{M}_X &\to \mathfrak{M}_X \\ (f, g) &\mapsto f \circ g\end{aligned}$$

It is associative but not communitative:
Let $f_a : x \mapsto a, f_b : x \mapsto b, \forall x \in X$ Then, $f_a \circ f_b = f_a, f_b \circ f_a = f_b$

Proposition 5.1.3 Let $(X, *)$ be an associative composition law on a set X.If $n \in \mathbb{N}_{>}0, x_1, \ldots, x_n \in X$, then , $\forall 1 \leq i \leq n-1$, we have

$$x_1 * \cdots * x_n = x_1 * \cdots * (x_i * x_{i+1}) * \cdots * x_n$$

Proof
$i = 1$: By definition, $x_1 * \cdots * x_n = (x_1 * x_2) * \cdots * x_n$.We suppose $i \geq 2$, by the associativity of $*$, we have

$$x_1 * \cdots * x_{i+1} = (x_1 * \cdots * x_{i-1}) * x_i * x_{i+1} = x_1 * \cdots * x_{i-1} * (x_i * x_{i+1})$$

$\square$

Definition 5.1.4 Let $(G, *)$ be a set equipped with a composition law, $g \in G$
If $\forall (x, y) \in G^2, g * x = g * y \Rightarrow x = y$, we say that g is **left cancellative**.
If $\forall (x, y) \in G^2, x * g = y * g \Rightarrow x = y$, we say that g is **right cancellative**.
If $*$ is communitative, left cancellative $\Leftrightarrow$ right cancellative.

Example 5.1.5
In $(\mathbb{N}, +)$, any element is cancellative.
In $(\mathbb{N}, *)$, any positive natural number is cancellative.

5.2 Neutral Element & Invertible Element

Definition 5.2.1 $(X, *), e \in X$ is called a **neutral element** if

$$\forall x \in X,\ e * x = x = x * e.$$

Proposition 5.2.2 Assume $(X, *)$ admits a neutral element, then its neutral element is unique.

Proof Let $e, e' \in X$ be neutral elements. Then

$$e = e * e' = e'.$$

$\square$

Definition 5.2.3 Let $(G, *)$ be a semigroup. If $(G, *)$ has a neutral element, then we say $(G, *)$ is **monoid**.

Example 5.2.4
(1) X is a set, $(\mathfrak{M}_x, \circ)$ is a monoid with the neutral element Id_X.
(2) $d \in \mathbb{N}_{>0}$, $(d\mathbb{N}, +)$ with neutral 0, $(\mathbb{N}, \times)$ with neutral 1.

Definition 5.2.5 Let $(G, *)$ be a monoid with the neutral element e. For any $(x, y) \in G^2$, if $x * y = e$ then we say x is a **left inverse** of y, and y is the **right inverse** of x.

Remark 5.2.6 We say x is **left invertible** if x has a left inverse. (resp. right invertible.)

Remark 5.2.7 x is left invertible in $(G, *) \Leftrightarrow x$ is right invertible in $(G, \hat{*})$.

Proposition 5.2.8 Let $(G, *)$ be a monoid, $g \in G$. If g is both left invertible and right invertible, then g has a unique left inverse and a unique right inverse, which actually coincide.

Proof Let x (resp. y) be a left (resp. right) inverse of g. Then , by the associativity law, we have

$$x = x * e = x * (g * y) = (x * g) * y = y.$$

Hence any left inverse is equal to y, hence it is unique. Similarly for the right. $\square$

Definition 5.2.9 Let $(G, *)$ be a monoid. If $g \in G$ is both left invertible and right invertible, then we say g is **invertible**. If g is invertible, the left inverse is equal to right inverse, hence we called it the inverse of g, denote by $\iota(g)$.

Proposition 5.2.10 Let $(G, *)$ be a monoid, $g \in G$. If g is right (resp. left) invertible, then it is right (resp. left) cancellative.

Proof Let h be the right inverse of g. If $x * g = y * g$, then

$$x = x * e = x * (g * h) = (x * g) * h = (y * g) * h = y * (g * h) = y * e = y.$$

$\square$

Notation 5.2.11 For a monoid $(G, *)$.
If $*$ is written multiplicatively, we usually denote $x * y$ as $x \cdot y$ or xy. If no ambiguity, neutral element as 1, inverse of x as x^{-1}.
If $*$ is written additively, $x * y$ as $x + y$, neutral element as 0, inverse of x as $-x$.

Proposition 5.2.12 Let $(G, *)$ be a monoid.
(1) If $x \in G$ is an invertible element , then $\iota(x)$ is also invertible, and $\iota(\iota(x)) = x$.
(2) If $x, y \in G$ are invertible, so does $x * y$ and $\iota(x * y) = \iota(y) * \iota(x)$.

Proof
(1)

$$x * \iota(x) = \iota(x) * x = e.$$

(2)

$$(xy)(\iota(y)\iota(x)) = xy\iota(y)\iota(x) = xe\iota(x) = x\iota(x) = e.$$

$$(\iota(y)\iota(x))(xy) = \iota(y)\iota(x)xy = \iota(y)ey = \iota(y)y = e.$$

$\square$

Definition 5.2.13 Let $(G, *)$ be a monoid. If any element of G is invertible, then we say G with the composition law is a **group**. A communitative group is also called **abelian group**.

Now we have :
(binary operations on X)$\supseteq$(semigroup)$\supseteq$(monoids)$\supseteq$(group)$\supseteq$(abelian group)

Example 5.2.14
(1) $(\mathbb{Z}, +)$ is an abelian group.
(2) Let X be a set and $\mathfrak{S}_X$ be the set of bijections from X to X.$(\mathfrak{S}_X, \circ)$ is a

monoid with the neutral element Id_X.Since $f \in \mathfrak{S}_X$ is bijective, hence there exists a unique inverse $f^{-1} \in \mathfrak{S}_X$.So $(\mathfrak{S}_x, \circ)$ is a group (but not abelian in general), called the symmetric group of X.
Let $\mathfrak{S}_n$ be the symmetric group of the set $\mathbb{N}_{\leq n}$, its element f can be denoted as a table:

$$\begin{pmatrix} 1 & 2 & \dots & n \\ f(1) & f(2) & \dots & f(n) \end{pmatrix}.$$

5.3 Substructure

Definition 5.3.1 Let $(G, *)$ be a semigroup, H be a subset of G. If H is close under $*$, then we say H is a **subsemigroup** of $(G, *)$. Note that H equipped with the restriction of $*$ forms a semigroup. Let $(G, *)$ be a monoid. If a sub-semigroup H of $(G, *)$ contains the neutral element of $(G, *)$, then we say H is a **submonoid** of $(G, *)$.

Example 5.3.2
(1) Let $d \in \mathbb{N}^*$, then $d\mathbb{N}$ forms a submonoid of $(\mathbb{N}, +)$. $d\mathbb{N}$ is a subsemigroup of $(\mathbb{N}, \cdot)$.
(2) $\mathfrak{S}_X$ is submonoid of $(\mathfrak{M}_X, \circ)$.

Proposition 5.3.3 Let $(M, *)$ be a monoid, $H \subseteq M$ be a non-empty subset. Suppose that any element of H is invertible in M , and $(\forall x, y \in H, (x, y) \mapsto x * \iota(y))$, if $\forall x, y \in H, x * \iota(y) \in H$, then H is a submonoid of M. Moreover, H equipped with the restriction of $*$ forms a group $(H, *|_H)$.

Proof Let e be the neutral element of $(M, *)$. Let $a \in H$, then $e = a * \iota(a) \in H$. For any $y \in H$, one has $\iota(y) = e * \iota(y) \in H$. For any $(x, y) \in H^2$, $x * y = x * \iota(\iota(y)) \in H$. Hence H is closed under $*$ and it contains the neutral element. Also, $\forall y \in H, \iota(y) \in H$, hence H is group. □

Corollary 5.3.4 Let $(M, *)$ be a monoid, G be the set of all invertible element in M. Then G is a submonoid. Moreover, G equipped with the restriction of $*$ forms a group.

Proof By definition, any element in G is invertible in M. By Proposition 5.2.12, $\forall x, y \in G, x * \iota(y) \in G$. Therefore, Proposition 5.3.3 implies the claim. □

Notation 5.3.5 Let M be a monoid, we often use $M^\times$ to denote the submonoid of M consisting of all invertible element if the composition law on M is not written additively.

Example 5.3.6 Let X be a set, $\mathfrak{M}_X^\times = \mathfrak{S}_X$.

Definition 5.3.7 Let $(G, *)$ be a group, $H \subseteq G$ be a submonoid. If $\forall x \in H$, one has $\iota(x) \in H$, then we say H is **subgroup** of G.

Proposition 5.3.8 Let $(M, *)$ be a monoid, $\varnothing \neq H \subseteq M^\times$ be a subset such that $\forall x, y \in H$,

$$x * \iota(y) \in H.$$

Then H is a subgroup of $M^\times$.

Proof Let e be the neutral element of $(M, *)$. By Proposition5.3.3, we obtain that H forms a submonoid of $M^\times$. Moreover, $\forall x \in H$, one has $\iota(x) = e * \iota(x) \in H$. So H is a subgroup of $M^\times$. □

Proposition 5.3.9 Let $(G, *)$ be a semigroup (resp. monoid, group), $(H_i)_{i \in I}$ be a family of subsemigroups (resp. submonoids, subgroups), where I is a non-empty set. Then

$$H := \bigcap_{i \in I} H_i$$

is a subsemigroup (resp. submonoid, subgroup) of G.

Proof
For semigroup case, let $x, y \in H$ then $x, y \in H_i, \forall i \in I$.Then $x * y \in H_i, \forall i \in I$, thus

$$x * y \in \bigcap_{i \in I} H_i = H.$$

For monoid case, the neutral element e of G satisfies

$$e \in H_i, \forall i \in I \Rightarrow e \in \bigcap_{i \in I} H_i = H.$$

For group case, to check $x * \iota(y) \in H$ like above. $\square$

5.4 Homomorphism

Definition 5.4.1 Let $(M, *)$ and $(N, \star)$ be semigroups, $f : M \to N$ be a mapping of sets.
(1) f is called a **semigroup homomorphism** from $(M, *)$ to $(N, \star)$ if

$$f(a * b) = f(a) \star f(b), \forall a, b \in M.$$

(2) If moreover, $(M, *)$ and $(N, \star)$ are both monoids with neutral elements e_M, e_N, f is called a **monoid homomorphism** if

$$f(a * b) = f(a) \star f(b), \forall a, b \in M,$$

$$f(e_M) = e_N.$$

(3) If moreover, $(M, *)$ and $(N, \star)$ are both groups, f is called a **group homomorphism** if

$$f(a * b) = f(a) \star f(b), \forall a, b \in M,$$

$$f(e_M) = e_N,$$

$$f(\iota(a)) = \iota(f(a)), \forall a \in M.$$

(They are not independent.)

Remark 5.4.2 Let $(M, *)$, $(N, \star)$ be groups, we claim that if $\forall a, b \in M, f(a * b) = f(a) \star f(b)$, then $f(e_M) = e_N$ and $f(\iota(a)) = \iota(f(a))$. Let $b = e_M$, then

$$f(e_M) = (\iota(f(a)) \star f(a)) \star f(e_M) = \iota(f(a)) \star (f(a) \star f(e_M)) = \iota(f(a)) \star f(a) = e_N.$$

$$\iota(f(a)) = \iota(f(a)) \star e_N = \iota(f(a)) \star (f(a) \star f(\iota(a))) = \iota(f(a)) \star f(a) \star f(\iota(a)) = f(\iota(a)).$$

But for monoid, we need $f(e_M) = e_N$.

Proposition 5.4.3
Let $f : (M, *) \to (N, \star)$ be a semigroup (resp. monoid, group) homomorphism. If M_1 is a subsemigroup (resp. submonoid, subgroup) of M, then the image $f(M_1)$ is a subsemigroup (resp. submonoid, subgroup).

Proof The semigroup case. Let $x, y \in f(M_1)$, we may write $x = f(a), y = f(b), a, b \in M_1$

$$x \star y = f(a) \star f(b) = f(a * b) \in f(M_1).$$

The monoid case.We denote e_M, e_N be the neutral elements of M, N

$$e_M \in M_1, e_N = f(e_M) \in f(M_1).$$

The group case.We have to check that $x, y \in f(M_1), x \star \iota(y) \in f(M_1)$

$$\forall a \in M, f(a) \star f(\iota(a)) = f(a * \iota(a)) = f(e_M) = e_N.$$

We may write $x = f(a), y = f(b), a, b \in M_1$

$$x \star \iota(y) = f(a) \star \iota(f(b)) = f(a) \star f(\iota(b)) = f(a * \iota(b)) \in f(M_1).$$

$\square$

Remark 5.4.4
(1) The semigroup homomorphism

$$f : (\mathbb{N}, \times) \to (\mathbb{N}, \times), n \mapsto 0$$

of two monoids, but is not a monoid homomorphism, and its image is $\{0\}$, which is not a submonoid of $(\mathbb{N}, \times)$.
(2) Let M be a semigroup (resp. monoid, group) and let N be a subsemigroup (resp. submonoid, subgroup). Then the inclusion mapping $\jmath : N \to M$ is a semigroup (resp. monoid, group) homomorphism.

Proposition 5.4.5 Let $(X, *) \xrightarrow{f} (Y, \star) \xrightarrow{g} (Z, \diamond)$ be semigroup (resp. monoid, group) homomorphisms. Then so does the composite mapping $g \circ f$.

Proof The semigroup case.

$$\begin{aligned}(g \circ f)(x_1 * x_2) &= g(f(x_1 * x_2)) = g(f(x_1) \star f(x_2)) \\ &= g(f(x_1)) \diamond g(f(x_2)), \forall x_1, x_2 \in X.\end{aligned}$$

The monoid case :

$$(g \circ f)(e_X) = g(f(e_X)) = g(e_Y) = e_Z.$$

The group case:

$$(g \circ f)(\iota(x)) = g(f(\iota(x))) = g(\iota(f(x))) = \iota((g \circ f)(x)).$$

$\square$

Proposition 5.4.6 Let $f : (X, *) \to (Y, \star)$ be a semigroup (resp.monoid, group) homomorphism between semigroups (resp.monoids groups). If f is bijective, then its inverse mapping $f^{-1} : Y \to X$ is also a semigroup homomorphism (resp.monoid, group)

Proof The semigroup case: Let $y_1, y_2 \in Y$ and let $x_i = f^{-1}(y_i)$, $i = 1, 2$. Then

$$y_1 \star y_2 = f(x_1) \star f(x_2) = f(x_1 * x_2),$$

$$f^{-1}(y_1 \star y_2) = x_1 * x_2 = f^{-1}(y_1) * f^{-1}(y_2).$$

The monoid case:

$$f(e_X) = e_Y \Rightarrow f^{-1}(e_Y) = e_X.$$

The group case:

$$f^{-1}(\iota(y)) \overset{y=f(x)}{=} f^{-1}(\iota(f(x))) = (f^{-1} \circ f)(\iota(x)) = \iota(f^{-1}(y)).$$

$\square$

Definition 5.4.7 A semigroup (resp. monoid, group) homomorphism $f : X \to Y$ is called a **semigroup (resp.monoid, group) isomorphism** if there exists a semigroup (resp.monoid, group) homomorphism $g : Y \to X$, such that

$$g \circ f = \mathrm{Id}_X, f \circ g = \mathrm{Id}_Y.$$

By Proposition 5.4.5, a semigroup (resp.monoid group) homomorphism is a semigroup (resp.monoid , group) isomorphism if and only if f is a bijection.

Proposition 5.4.8 Let $(G, *)$ be a group. The inversion mapping $\imath : (G, *) \to (G, \hat{*})$ is a group isomorphism.

5.5 Quotient

Definition 5.5.1 Let X be a set and $\sim$ be a binary relation on X. (We write $x \sim y$ the condition $(x, y) \in \Gamma_{\sim}$)
(1) If $\forall x \in X, x \sim x$.
(2) $\forall (x, y) \in X^2, x \sim y \Rightarrow y \sim x$.
(3) $\forall (x, y, z) \in X^3, (x \sim y$ and $y \sim z) \Rightarrow x \sim z$.
We say that $\sim$ is a **equivalence relation**.

Check Section 4.2:Equivalent Relation, to get more information about it.

Proposition 5.5.2 Let $(X_i)_{i\in i}$ be a family of sets. For any $i \in I$, let $\sim_i$ be an equivalence relation on X_i. Let $X = \prod_{i\in I} X_i$. We define a binary relation $\sim$ on X as follows:

$$(x_i)_{i\in I} \sim (y_i)_{i\in I} \Leftrightarrow \forall i \in I, x_i \sim_i y_i.$$

Then, $\sim$ is an equivalence relation, and the mapping

$$X/\sim \xrightarrow{\Phi} \prod_{i\in I} X_i/\sim_i,$$

$$[(x_i)_{i\in I}] \longmapsto ([x_i])_{i\in I}$$

is a bijection.

Proof
(1) Let $(x_i)_{i\in I} \in X. \forall i \in I, x_i \sim x_i$, so $(x_i)_{i\in I} \sim (x_i)_{i\in I}$.
(2) Let $x = (x_i)_{i\in I}, y = (y_i)_{i\in I}, x_i \sim_i y_i$, so $y_i \sim x_i$. Therefore, $y \sim x$.
(3) Let $x = (x_i)_{i\in I}, y = (y_i)_{i\in I}, z = (z_i)_{i\in I}$ in X. If $x \sim y$ and $y \sim z$, then $\forall i \in I, x_i \sim_i y_i$ and $y_i \sim_i z_i$. Hence $\forall i \in I, x_i \sim_i z_i$. So $x \sim z$.
We check that Φ is well defined. Let $x = (x_i)_{i\in I}$ and $y = (y_i)_{i\in I}$ be elements of X. If $[x] = [y]$, then $x \sim y$ and hence $\forall i \in I, x_i \sim_i y_i$, that means

$$([x_i])_{i\in I} = ([y_i])_{i\in I}.$$

By definition, Φ is surjective. If $\Phi([(x_i)_{i\in I}]) = \Phi([(y_i)_{i\in I}])$, then $\forall i \in I, [x_i] = [y_i]$, namely $x_i \sim_i y_i$. Therefore, $([(x_i)_{i\in I}]) = ([(y_i)_{i\in I}])$. $\square$

Notation 5.5.3 Let X be a set , $\sim$ be an equivalence relation on X. Then $X/\sim$

is called the **quotient** of X by $\sim$. The mapping

$$\pi : X \longrightarrow X/\sim,$$

$$x \longmapsto [x]$$

is called the **quotient mapping**.

Definition 5.5.4 Let X be a set, $f : X \to Y$ be a mapping and $\sim$ an equivalence relation on X. If $\forall (x, y) \in X^2, x \sim y \Rightarrow f(x) = f(y)$ we say that $\sim$ is **compatible** with f.

Theorem 5.5.5 (Proposition4.2.5) Let $f : X \to Y$ be a mapping and $\sim$ be an equivalence relation on X which is compatible with f. Then there exists a unique mapping

$$\tilde{f} : X/\sim \to Y, [x] \mapsto f(x),$$

such that

$$\tilde{f} \circ \pi = f.$$

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[d, "\pi"'] & Y \\
X/\sim \arrow[ur, "\tilde{f}"'] &
\end{tikzcd}$$

Proof If such $\tilde{f}$ exists. For an $x \in X$.

$$\tilde{f}([x]) = \tilde{f}(\pi(x)) = f(x)$$

So $\tilde{f}$ is unique. To prove the existence, it suffices to check that $\tilde{f} : X/\sim \to Y$ is well defined. If $[x] = [y]$, then $[x] \mapsto f(x), x \sim y$ and hence $f(x) = f(y)$. So $\tilde{f}$ is well defined. $\square$

Definition 5.5.6 We call $\tilde{f}$ the **mapping induced by** f **by passing to quotient**.

Example 5.5.7 Let X be a set and $*$ a composition law on X. We say that an equivalence relation $\sim$ on X is compatible to $*$ if $\forall (x_1, y_1), (x_2, y_2) \in X^2$

$$(x_1 \sim x_2 \text{ and } y_1 \sim y_2) \Leftrightarrow x_1 * y_1 \sim x_2 * y_2$$

Or equivalently, the equivalence relation R on $X \times X$ defined by

$$(x_1, y_1)R(x_2, y_2) \Leftrightarrow x_1 \sim x_2 \text{ and } y_1 \sim y_2$$

is compatible with the mapping:

$$X \times X \longrightarrow X/\sim$$

$$(x, y) \longmapsto [x * y]$$

By the theorem, $*$ induces by passing to quotient a mapping

$$(X/\sim) \times (X/\sim) \longrightarrow (X \times X)/R \longrightarrow X/\sim$$

$$([x], [y]) \longmapsto [(x, y)] \longmapsto [x * y]$$

The compatible mapping

$$(X/\sim) \times (X/\sim) \longrightarrow X/\sim$$

$$([x], [y]) \longmapsto [x * y]$$

defines a composition law on $X/\sim$, which is often denoted as $*$ by abuse of notation, called the composition law on $X/\sim$ induced by the composition law $*$ on X by passing to quotient.

Example 5.5.8 N_n on $\mathbb{Z}$.
If $n \mid (x_1 - x_2)$ $n \mid (y_1 - y_2)$, then $n \mid (x_1 + y_1) - (x_2 + y_2)$.
Since $x_1y_1 - x_2y_2 = x_1(y_1 - y_2) + (x_1 - x_2)y_2$, $n \mid x_1y_1 - x_2y_2$.
Hence $+$ and $\cdot$ on $\mathbb{Z}$ induces by passing to equivalent composition law on $\mathbb{Z}/\sim_n$.

Proposition 5.5.9
(1) If $* : X \times X \to X$ is associative (resp.communitative) then so is

$$* : (X/\sim) \times (X/\sim) \to X/\sim$$

(2) If e is a neutral element of $(X, *)$, then $[e]$ is a neutral element of $(X/\sim, *)$.
(3) If $(X, *)$ is a semigroup (resp. monoid), then the projection

$$\pi : X \to X/\sim, x \mapsto [x]$$

is a homomorphism of semigroup (resp.monoid).
(4) If $(X, *)$ is a monoid, $x \in X$ is invertible, then $[x]$ is invertible in $(X/\sim, *)$.

Proof
(1)associative: $[x] * ([y] * [z]) = [x] * [y * z] = [x * (y * z)] = [(x * y) * z] = [x * y] * [z] = ([x] * [y]) * [z]$.
communitative: $[x] * [y] = [x * y] = [y * x] = [y] * [x]$
(2) $[e] * [x] = [e * x] = [x], [x] * [e] = [x * e] = [x]$.
(3)
$$\pi(x * y) = [x * y] = [x] * [y] = \pi(x) * \pi(y)$$
$\pi(e) = [e]$ is the neutral element of $(X/\sim, *)$.
(4) By (3), π is a homomorphism of monoid, $\forall x \in X^{\times}, \pi(x) = [x] \in (X/\sim)^{\times}$ and $\iota([x]) = [\iota(x)]$. $\square$

Remark 5.5.10 If $(X, *)$ is a group, so is $(X/\sim, *)$.

Definition 5.5.11
If $(X, *)$ is a semigroup (resp. monoid, group), then $(X/\sim, *)$ is called the **quotient semigroup** (resp. quotient monoid, quotient group) of $(X, *)$ by $\sim$.

Definition 5.5.12 Let $(X, *), (Y, \star)$ be groups and $f : X \to Y$ be a homomorphism of groups. We define the **kernel** of f as

$$\ker(f) := \{x \in X \mid f(x) = e_Y\}$$

where e_Y is the neutral element of Y.

Proposition 5.5.13 Let $(X, *)$ be a monoid, $(Y, \star)$ be a semigroup, $f : X \to Y$ be a homomorphism of semigroups. If f is surjective, then $(Y, \star)$ is a monoid, and f is a homomorphism of monoid.

Proof We check that $f(e_X)$ is the neutral element of Y. $\forall y \in Y, \exists x \in X, f(x) = y$. So $f(e_X) \star y = f(e_X) \star f(x) = f(e_X * x) = f(x) = y$. Also $y \star f(e_X) = f(x) \star f(e_X) = f(x * e_X) = f(x) = y$. $\square$

Proposition 5.5.14 Let $(X, *)$ be a monoid and $(Y, \star)$ be a group. If $f : X \to Y$ is a homomorphism of semigroups, then it is homomorphism of monoids.

Proof Let e_X and e_Y be neutral elements of X and Y. One has $e_X = e_X * e_X$, so $f(e_X) = f(e_X) \star f(e_X)$, so $e_Y = f(e_X)$. $\square$

Proposition 5.5.15
(1) $\ker(f)$ is a subgroup of X.
(2) $\forall(a, x) \in X \times \ker(f)$, there exists $y \in \ker(f)$ such that $a * x = y * a$.

Proof
(1) The neutral element e_x of $(X, *)$ belongs to $\ker(f)$. If x, y are elements of $\ker(f)$, then

$$f(x * \iota(y)) = f(x) \star f(\iota(y)) = f(x) \star \iota(f(y)) = e_Y \star \iota(e_Y) = e_Y,$$

so, $x * \iota(y) \in \ker(f)$.
(2) We should take $y := (a * x) * \iota(a)$. It remains to check that $y \in \ker(f)$. One has $f(y) = f(a * x * \iota(a)) = f(a) \star f(x) \star \iota(f(a)) = f(a) \star \iota(f(a)) = e_Y$. $\square$

Definition 5.5.16 Let $(G, *)$ be a group and H be a subgroup of G. If $\forall(a, x) \in G \times H$, $a * x * a^{-1} \in H$, we say that H is a **normal subgroup**.

Proposition 5.5.17 Let $(G, *)$ be a group and H be a normal subgroup of G.
(1) The binary relation $\sim_H$ on G defined as

$$x \sim_H y \Leftrightarrow x * \iota(y) \in H$$

is an equivalence relation on G. Moreover,

$$\forall x \in G, [x] = H * x := \{y * x \mid y \in H\}.$$

(2) If H is normal, then

$$\forall x \in G, x * H = H * x.$$

Moreover, $\sim_H$ is compatible with $*$.
(3) The kernel of $\pi : G \to G/\sim_H$ is equal to H.

Proof
(1) If $x \sim_H y$, then $x * \iota(y) \in H$, so $y * \iota(x) = \iota(x * \iota(y)) \in H$, so $y \sim_H x$. If $x \sim_H y$ and $y \sim_H z$, then $x * \iota(y) \in H, y * \iota(z) \in H$, so $x * \iota(z) = x * \iota(y) * y * \iota(z) \in H$. Hence $x \sim_H z$. By definition, $[x] := \{y \in G \mid x * \iota(y) \in H\}$. If $y \in [x]$, then $y * \iota(x) \in H$. Hence $y = (y * \iota(x)) * x \in H * x$. Conversely, if $y = h * x \in H * x \quad (h \in H)$, then $y * \iota(x) = h * x * \iota(x) \in H$. So $y \in [x]$,

$[x] = H * x$.
We denote by G/H the set

$$G/H := \{x * H \mid x \in G\}.$$

We denote by $H\backslash G$ the set

$$H\backslash G := \{H * x \mid x \in G\}.$$

(2) Suppose that H is normal.$\forall (x, y) \in G \times H$, one has $x * y * \iota(x) \in H$. So $\forall y \in H, \exists z (= x * y * \iota(x)) \in H$ such that $x * y = z * x$. So $x * H \subseteq H * x$. Conversely, $H * x \subseteq x * H$. Let x_1, x_2, y_1, y_2 be elements of G, such that $x_1 \sim_H x_2, y_1 \sim_H y_2$.

$$\begin{aligned}&(x_1 * y_1) * \iota(x_2 * y_2)\\ =&x_1 * y_1 * \iota(y_2) * \iota(x_2)\\ =&x_1 * (y_1 * \iota(y_2)) * \iota(x_1) * x_1 * \iota(x_2) \in H.\end{aligned}$$

(3)
$$\ker(\pi) = [e_G] = H * e_G = H.$$

□

Notation 5.5.18 If H is a normal subgroup of G, we denote by G/H the quotient group $G/\sim_H$.

Theorem 5.5.19 Let $f : (X, *) \to (Y, \star)$ be a homomorphism of groups, and $K = \ker(f)$. Then $\sim_K$ is compatible with f, and f induces by passing to quotient a mapping

$$\tilde{f} : X/K \longrightarrow Y,$$

which is actually an injective homomorphism of groups, with $\tilde{f}(X/K) = f(X)$. In particular, X/K is isomorphism to $f(X)$.

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "\pi"'] & f(X) \subseteq Y \\ X/\ker(f) \arrow[ur, "\tilde{f}"'] & \end{tikzcd}$$

Proof Let x and y be elements of X. $x \sim_K y \Leftrightarrow x * \iota(y) \in K$. Hence $f(x) \star \iota(f(y)) = f(x * \iota(y)) = e_Y$. So $f(x) = f(y)$. $\tilde{f}([x] * [y]) = \tilde{f}([x * y]) = f(x * y) = f(x) \star f(y) = \tilde{f}([x]) \star \tilde{f}([y])$. $\square$

5.6 Universal Homomorphisms

Proposition 5.6.1 Let $(M, *)$ be a monoid, $x \in M$. Then there exists a unique homomorphism of monoid $f : (\mathbb{N}, +) \to (M, *)$ such that $f(1) = x$.

Proof We construct a mapping $f : \mathbb{N} \to M$ in a recursive way as follows: $f(0) = e_M$. For any $n \in \mathbb{N}$, we let $f(n+1) = f(n) * x$. We will prove that f is a homomorphism of monoids, that is

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

We reason by induction on m. If $m = 0$, $f(n) = f(n) * e_M$. Suppose that $f(n+m) = f(n) * f(m)$. One has

$$f(n+m+1) = f(n+1) * f(m) = f(n) * f(1) * f(m) = f(n) * f(m+1).$$

If $g : \mathbb{N} \to M$ is a homomorphism of monoid, such that $g(1) = x$. Since $g(n+1) = g(n) * g(1) = g(n) * x$, we have $g(n) = f(n)$. By induction, $\forall n \in \mathbb{N}, g(n) = f(n)$. So f is unique. $\square$

Notation 5.6.2 Let $(M, *)$ be a monoid, $x \in M$, $f : (\mathbb{N}, +) \to (M, *)$ be the unique homomorphism of monoid, such that $f(1) = x$. For any $n \in \mathbb{N}$, we denote by x^{*n} the element $f(n) \in M$, $x^{*0} = e_M$, $x^{*(n+m)} = x^{*n} * x^{*m}$.
Two exceptions: If $* = \cdot$ is written multiplicatively, x^{*n} is written as x^n. If $* = +$, then x^{*n} is written as nx.

Proposition 5.6.3 Let $(M, *)$ be a monoid, $x \in M$. There exists a unique homomorphism of monoids $f : (\mathbb{Z}, +) \to (M, *)$ such that $f(1) = x$. Note that $f(\mathbb{Z}) \subseteq M^\times$. So f defines a homomorphism of groups $f : (\mathbb{Z}^\times, +) \to (M^\times, *)$.

Proof We define f as

$$f(n) := \begin{cases} x^{*n}, & n \geq 0 \\ \iota(x^{*(-n)}), & n < 0 \end{cases}.$$

Let n, m be two elements of $\mathbb{Z}$.
(1) If $n, m > 0$. Then $f(n+m) = x^{*n} * x^{*m} = f(n+m)$.
(2) If $n, m < 0$. Then $f(n+m) = \iota(x^{*(-n-m)}) = \iota(x^{*(-m)} * x^{*(-n)}) = \iota(x^{*(-n)}) * \iota(x^{*(-m)}) = f(n) * f(m)$.
(3) If $n > 0, m < 0$ and $n + m > 0$. Then

$$f(n+m) = x^{*(n-(-m))} = x^{*n} * \iota(x^{*(-m)}) = f(n) * f(m).$$

(4) If $n > 0, m < 0$ and $n + m < 0$. Then

$$f(n+m) = \iota(x^{*(-n-m)}) = \iota(\iota(x^{*n}) * x^{*(-m)}) = \iota(x^{*(-m)}) * x^{*n} = f(m) * f(n).$$

□

Notation 5.6.4 If $x \in M^{\times}$, for any $n \in \mathbb{Z}$, let x^{*n} be the image of n by this unique homomorphism of monoids $(\mathbb{Z}, +) \to (M, *)$, $1 \mapsto x$. $x^{\cdot n}$ is denoted as x^n, x^{+n} is denoted as nx.

Proposition 5.6.5 Let $(M, *)$ be a monoid, $x, y \in M$.
(1) If $x * y = y * x$, then for any $(n, m) \in \mathbb{N}^2$,

$$x^{*n} * y^{*m} = y^{*m} * x^{*n}.$$

$$(x * y)^{*n} = x^{*n} * y^{*n}.$$

(2) If $x \in M$, $\iota(x^{*n}) = \iota(x)^{*n}$ and for any $(n, m) \in \mathbb{N}^2$, with $n \geq m$,

$$x^{*(n-m)} = x^{*n} * \iota(x)^{*m},$$

$$\iota(x^{*(n-m)}) = \iota(x)^{*n} * x^{*m}.$$

Proof
(1) We prove by induction on n such that $x^{*n} * y = y * x^{*n}$. If $n = 0$, $x^{*n} = e_M$, so $y * e_M = y = e_M * y$. If $x^{*n} * y = y * x^{*n}$, we have $x^{*n+1} * y = x^{*n} * y * x = y * x^{*n} * x = y * x^{*(n+1)}$. We apply this statement in replacing n by m, x by y, and y by x^{*n}. From $x^{*n} * y = y * x^{*n}$, we deduce that $y^{*m} * x^{*n} = x^{*n} * y^{*m}$. We prove $(x * y)^{*n} = x^{*n} * y^{*n}$ by induction on n. If $n = 0, e_M = e_M * e_M$. If $n = 1, x * y = x * y$. If $(x * y)^{*n} = x^{*n} * y^{*n}$, then

$$(x*y)^{*(n+1)} = (x*y)^{*n}*x*y = x^{*n}*y^{*n}*x*y = x^{*n}*x*y^{*n}*y = x^{*(n+1)}*y^{*(n+1)}.$$

(2) $x^{*n} * \iota(x)^{*n} = (x * \iota(x))^{*n} = e_M^{*n} = e_M$, since $(\mathbb{N}, +) \to (M, *), n \mapsto e_M$ is a homomorphism of monoids. $\iota(x)^n * x^{*n} = (\iota(x) * x)^{*n} = e_M$. If $n \geq m$

$$x^{*n} * \iota(x)^{*m} = x^{*(n-m)} * x^{*m} * \iota(x)^{*m} = x^{*(n-m)}.$$

$$\iota(x)^{*n} * x^{*m} = \iota(x)^{*(n-m)} * \iota(x)^{*m} * x^{*m} = \iota(x)^{*(n-m)} = \iota(x^{*(n-m)}).$$

□

Definition 5.6.6 Let I be a set. For any $i \in I$, let $(M_i, *_i)$ be a set equipped with a composition law. Let

$$M = \prod_{i \in I} M_i = \{(x_i)_{i \in I} \mid \forall i \in I, x_i \in M_i\}.$$

We define a composition law on M such that

$$(x_i)_{i \in I} * (y_i)_{i \in I} = (x_i * y_i)_{i \in I}.$$

For any $j \in I$, let $\pi_j : M \to M_j,\ (x_i)_{i \in I} \mapsto x_j$.

Proposition 5.6.7
(1) If $\forall i \in I, *_i$ is commutative, then $*$ is communitative.
(2) If $\forall i \in I, *_i$ is associative, then $*$ is associative. Moreover, $\pi_j : (M, *) \to (M_j, *_j)$ is a homomorphism of semigroups.
(3) If $\forall i \in I$, e_i is a neutral element of $(M_i, *_i)$, then $e := (e_i)_{i \in I}$ is a neutral element of $(M, *)$. Moreover, if each $(M_i, *_i)$ is a monoid, then $\pi_j : (M, *) \to (M_j, *_j)$ is a homomorphism of monoids.
(4) Assume that each $(M_i, *_i)$ is a monoid. Then $M^\times = \prod_{i \in I} M_i^\times$. In particular, if each $(M_i, *_i)$ is a group, then $M^\times = \prod_{i \in I} M_i^\times$ is also a group.

Proof If $(x_i)_{i \in I}, (y_i)_{i \in I} \in M$, then $\pi_j(x * y) = \pi_j((x_i * y_i)_{i \in I}) = x_j *_j y_j = \pi_j(x) * \pi_j(y)$.
proof of (4): Assume that $x = (x_i)_{i \in I} \in M^\times$. Then $\exists y = (y_i)_{i \in I} \in M^\times$ such that $x * y = e := (e_i)_{i \in I}$, where e_i is the neutral element of $(M_i, *_i)$. $x * y = (x_i * y_i)_{i \in I} = (e_i)_{i \in I} =$. So $x_i *_i y_i = e_i$ for all $i \in I$. Therefore, $x_i \in M_i^\times$ for all $i \in I$. Hence $M^\times \subseteq \prod_{i \in I} M_i^\times$. Now let $x = (x_i)_{i \in I} \in \prod_{i \in I} M_i^\times$. We claim that $(\iota(x_i))_{i \in I}$ is the inverse of x. In fact $(x_i)_{i \in I} * (\iota(x_i))_{i \in I} = e$. So $x \in M^\times$. □

Theorem 5.6.8 Suppose that each $(M_i, *_i)$ is a semigroup. Let $(N, \star)$ be a semigroup (resp. monoid, group). For any $i \in I$, let $f_i : N \to M_i$ be a homomorphism of semigroups (resp. monoid, group). Then there is a unique homomorphism of semigroups (resp. monoid, group) $f : M \to N$ such that $\forall i \in I, \pi_i \circ f = f_i$. $(M, *)$ is called the **product** of $(M_i, *_i)$.

Proof By Proposition 3.9.5, there exists a unique mapping $f : N \to M$ such that $\forall i \in I, \pi_i \circ f = f_i$. We check that f is a homomorphism.
Recall that $\forall y \in N, f(y) = (f_i(y))_{i\in I}$. If $(y, z) \in N \times N$, then $f(y * z) = (f_i(y) * f_i(z))_{i\in I} = (f_i(y))_{i\in I} * (f_i(z))_{i\in I} = f(y) * f(z)$. If each $(M_i, *_i)$ is a monoid with neutral element e_i, and e_N is the neutral element of N, in the case where each f_i is a homomorphism of monoids $(f_i(e_N) = e_i)$. One has $f(e_n) = (f_i(e_N))_{i\in I} = (e_i)_{i\in I}$ is the neutral element of M. □

Notation 5.6.9 Let M be a communitative monoid, $(x_i)_{i\in I}$ be a family of elements in M. We suppose that $I_0 = \{i \in I \mid x_i \neq e\}$ is finite. We pick a natural number n and a bijection $\sigma : \{1, 2, \ldots, n\} \to I_0$. If the composition law of M is written as $+$, then

$$\sum_{i\in I} x_i \text{ denotes } (x_{\sigma(1)} + x_{\sigma(2)} + \cdots + x_{\sigma(n)}),$$

it denotes the neutral element 0 of M when $I_0 = \varnothing$. If the composition law of M is written as $\cdot$, then

$$\prod_{i\in I} x_i \text{ denotes } (x_{\sigma(1)} \cdot x_{\sigma(2)} \cdots x_{\sigma(n)}),$$

it denotes the neutral element 1 of M when $I_0 = \varnothing$.
Let $(M_i)_{i\in I}$ be a family of communitative monoids. (The composition law of M_i is written additively, the neutral element of M_i is written 0)

Notation 5.6.10 Let $(M_i)_{i\in I}$ be a family of communitative monoids. For any $i \in I$, let e_i be a neutral element of M_i. We denote by

$$\bigoplus_{i\in I} M_i$$

the set of $(x_i)_{i\in I} \in \prod_{i\in I} M_i$ such that $\{i \in I \mid x_i \neq e_i\}$ is finite.

Proposition 5.6.11 $\bigoplus_{i\in I} M_i$ is a submonoid of $\prod_{i\in I} M_i$.

Proof First, $e := (e_i)_{i\in I} \in \bigoplus_{i\in I} M_i$. Let $*_i$ be the composition law of M_i, $*$ be the direct product of $(*_i)_{i\in I}$. If $x = (x_i)_{i\in I}$ and $y = (y_i)_{i\in I}$ are in $\bigoplus_{i\in I} M_i$, then $x * y = (x_i * y_i)_{i\in I}$. If $I_x = \{i \in I \mid x_i \neq e_i\}$ and $I_y = \{i \in I \mid y_i \neq e_i\}$ are finite, then $\{i \in I \mid x_i * y_i \neq e\} \subseteq I_x \cup I_y$. So $x \in \bigoplus_{i\in I} M_i$ and $y \in \bigoplus_{i\in I} M_i$ imply that $x * y \in \bigoplus_{i\in I} M_i$. $\square$

Definition 5.6.12 (Direct sum) $\bigoplus_{i\in I} M_i$ is called the **direct sum** of $(M_i)_{i\in I}$. For any $j \in I$, the homomorphisms

$$\begin{aligned} M_j &\xrightarrow{\mathrm{Id}_{M_j}} M_j \\ M_j &\longrightarrow M_i, \quad (i \neq j) \\ x_j &\longmapsto e_i \end{aligned}$$

induce:

$$\begin{aligned} M_j &\longrightarrow \textstyle\prod_{i\in I} M_i \\ x_j &\longmapsto (y_i)_{i\in I} \end{aligned}$$

with

$$y_i = \begin{cases} x_j, & j = i \\ e_i, & i \neq j \end{cases}.$$

Claim: This homomorphism takes value in $\bigoplus_{i\in I} M_i$. We denote by

$$\lambda_j : M_j \longrightarrow \bigoplus_{i\in I} M_i$$

this homomorphism.

$$\lambda_j(x_j)_i = \begin{cases} x_j, & i = j \\ e_i, & i \neq j \end{cases},$$

$$\lambda_j(x_j)_i = (\lambda_j(x_j)_i)_{i\in I}.$$

Theorem 5.6.13 Let $(N, \star)$ be a communitative monoid. Then for any $i \in I$, let $\psi_i : M_i \to N$ be a homomorphism of monoids. Then there is a unique homomorphism of monoids $\psi : \bigoplus_{i\in I} M_i \to N$ such that for any $j \in I$, $\psi \circ \lambda_j = \psi_j$.

$$\begin{tikzcd}
\bigoplus_{i\in I} M_i \arrow[r, "\psi"] & N \\
M_j \arrow[u, "\lambda_j"] \arrow[ur, "\psi_j"'] &
\end{tikzcd}$$

Proof For simplicity, we write all composition laws as $+$, and all neutral element as 0. We should define $\psi : \bigoplus_{i\in I} M_i \to N,\ (x_i)_{i\in I} \mapsto \sum_{i\in I} \psi_i(x_i)$. $\psi\left((0)_{i\in I}\right) = \sum_{i\in I} 0 = 0$. $\psi\left((x_i)_{i\in I} + (y_i)_{i\in I}\right) = \psi\left((x_i + y_i)_{i\in I}\right) = \sum_{i\in I} \psi_i(x_i + y_i) = \sum_{i\in I} [\psi_i(x_i) + \psi_i(y_i)] = \sum_{i\in I} \psi_i(x_i) + \sum_{i\in I} \psi_i(y_i)$. (The last equality holds because the composition law of N is commutative.) $\square$

Chapter 6

Rings and Modules

6.1 Unitary Rings

Definition 6.1.1 Let A be a set, and $+$ and $*$ be composition laws. If
(1) $(A, +)$ forms a communitative group.
(2) $(A, *)$ forms a monoid.
(3) For any $(a, b, c) \in A^3$, $a*(b+c) = (a*b)+(a*c)$ and $(b+c)*a = (b*a)+(c*a)$.
(4)† If in addition, $*$ is communitative, then we say that the unitary ring $(A, +, *)$ is communitative.

Example 6.1.2 $(\mathbb{Z}, +, \cdot)$ is a unitary ring.

Note that, if we denote by $\hat{*}$ the composition law

$$A \times A \longrightarrow A,$$

$$(a, b) \longmapsto b * a.$$

Then $(A, +, \hat{*})$ forms a unitary ring. We call it the opposite unitary ring of $(A, +, *)$.

Notation 6.1.3 Usually, we denote by $+$ the first composition law, of a unitary ring A and call it the **addition**. We denote by 0 the neutral element of $+$, and call it the **zero element** of A. Usually we denote by $\cdot$ the second composition law of A and call it the **multiplication**. We denote by 1 the neutral element with respect to $\cdot$, and call it the **unity element** of A.

Definition 6.1.4 Let A be a unitary ring and B be a subset of A. If B is a subgroup of $(A, +)$ and a submonoid of $(A, \cdot)$, then we call B a **unitary subring** of A.

Example 6.1.5 Let $\{0\}$ be the set of 1 element. Let $+$ and $\cdot$ be both the composition law $\{0\} \times \{0\} \to \{0\}, (0,0) \mapsto 0$. Then $(\{0\}, +, *)$ is a unitary ring. We call it the **zero ring**.

Definition 6.1.6 Let A and B be unitary rings and $f : A \to B$ be a mapping. If f is a group homomorphism from $(A, +)$ to $(B, +)$, and is a monoid homomorphism from $(A, \cdot)$ to $(B, \cdot)$, then we call f a **unitary ring homomorphism**.

Proposition 6.1.7 For any unitary ring A, there exists a unitary ring homomorphism $A \to \{0\}$.

Lemma 6.1.8 Let A be a unitary ring.
(1) $\forall a \in A, 0a = a0 = 0$.
(2) $\forall a, b \in A, -(ab) = (-a)b = a(-b)$.

Proof
(1) $0 + 0 = 0$, so $0 + 0a = 0a = (0 + 0)a = 0a + 0a$. Hence $0a = 0$.
(2) $ab + (-a)b = (a + (-a))b = 0b = 0, ab + a(-b) = a(b + (-b)) = a0 = 0$.
□

Proposition 6.1.9 For any unitary ring A, there exists a unitary ring homomorphism from $\mathbb{Z}$ to A.

Proof If $f : \mathbb{Z} \to A$ is a unitary ring homomorphism, then $f(1) = 1_A$. So f is identifies with the unitary group homomorphism.

$$(\mathbb{Z}, +) \longrightarrow (A, +),$$

$$n \longmapsto n1_A.$$

It remains to check that for any $(n, m) \in \mathbb{Z}^2, f(nm) = f(n)f(m)$. Note that , if $(n, m) \in \mathbb{N} \times \mathbb{N}$, then

$$f(n) = \underbrace{1_A + \cdots + 1_A}_{n \text{ copies}}, \ f(m) = \underbrace{1_A + \cdots + 1_A}_{m \text{ copies}}.$$

So $f(n)f(m) = nm1_A1_A = nm1_A = f(nm)$. $f(-n)f(m) = (-f(n))f(m) = -f(n)f(m) = -f(nm) = f(-nm)$. $f(-n)f(-m) = \ldots$ □

Definition 6.1.10 Let K be a unitary ring. We denote by $K^\times$ the invertible elements of $(K, \cdot)$. If $K^\times = K\backslash\{0\}$ then we say that K is a division ring. If in addition, K is commutative, then we say that K is a **field**.

Example 6.1.11 $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are fields.

6.2 Action of Monoids

Definition 6.2.1 Let $(G, *)$ be a monoid, the neutral element of which is denoted as e. Let X be a set. We call **left action** of G on X any mapping

$$\phi : G \times X \to X,$$

such that
(1) $\phi(e, x) = x$, for any $x \in X$.
(2) $\forall (a, b) \in G \times G, \forall x \in X$,

$$\phi(a * b, x) = \phi(a, \phi(b, x)).$$

(Resp. right action)

Remark 6.2.2 A left action of $(G, *)$ on X is a right action of $(G, \hat{*})$ on X.

Notation 6.2.3 If $* = \cdot$, a left action is usually denoted as

$$G \times X \longrightarrow X,$$

$$(a, x) \longmapsto ax.$$

Condition (1) becomes $ex = x$, (2) becomes $(ab)x = a(bx)$.

Example 6.2.4 Let G be a group, H be a subgroup of G. Then

$$H \times G \longrightarrow G,$$

$$(h, g) \longmapsto hg.$$

is a left action of H on G. (Resp. right action.)

Proposition 6.2.5 Let G be a monoid, X be a set and $\phi : G \times X \longrightarrow X$ be a left action of G on X. We define a binary relation $\sim_\phi$ on X as follows:

$$x \sim_\phi y \Leftrightarrow \exists g \in G, \phi(g, x) = y.$$

Then $\sim_\phi$ is reflexive and transitive. It is an equivalence relation if G is a group.

Proof
Reflexivity: Let e be the neutral element of G, then $x = ex$, so $x \sim_\phi x$.
Transitivity: If $y = ax$ and $z = by$, then $z = b(ax) = (ba)x$, so $x \sim_\phi y \wedge y \sim_\phi z \Rightarrow x \sim_\phi z$.
Assume that G is a group. If $y = ax$, then $\iota(a)y = \iota(a)(ax) = (\iota(a)a)x = ex = x$, so $x \sim_\phi y$ implies $y \sim_\phi x$. $\square$

Definition 6.2.6 Let G be a group, X be a set and $\phi : G \times X \longrightarrow X$ be a left action. For any $x \in X$, the equivalence class of x under the equivalence relation $\sim_\phi$ is called the **orbit** of x under the action ϕ, denoted as $\mathrm{orb}_\phi(x)$. We denote by $G\backslash X$ the set of all orbits of X under the action ϕ. (Resp. right action and X/G.)

Remark 6.2.7 If X is finite, then

$$\mathrm{card}(X) = \sum_{A \in G\backslash X} \mathrm{card}(A).$$

In particular, if $(G, *)$ is a finite group, and H is a subgroup of G, then $\mathrm{card}(G) = \mathrm{card}(H)\mathrm{card}(H\backslash G)$. In fact, $H\backslash G = \{H{*}x \mid x \in G\}$, $H{*}x := \{h{*}x \mid h \in H\}$.

6.3 Vector Space

Definition 6.3.1 Let K be a unitary ring. Let $(V, +)$ be an abelian group. (Neutral element of $(V, +)$ is denote as 0.) We call a **left K-module structure** any left action of $(K, \cdot)$ on V.

$$\phi : K \times V \longrightarrow V$$

(1) $\forall (a, b) \in K \times K, \forall x \in V$,

$$\phi(a + b, x) = \phi(a, x) + \phi(b, x).$$

(2) $\forall a \in K, \forall (x,y) \in V \times V$,

$$\phi(a, x+y) = \phi(a,x) + \phi(a,y).$$

The abelian group $(V,+)$ equipped with a left K-module structure is called a **left K-module**. If K is communitative, left and right K-modules structures have the same axioms. So we just call them K-module structures. Left and right K-modules structures are called K-modules. If K is a field, a K-module is called a **vector space** over K.

Example 6.3.2 $(\{0\},+)$ is a left K-module. Action

$$\phi : K \times \{0\} \longrightarrow \{0\},$$

$$\phi(a,0) = 0.$$

It is called the zero K-module.

Example 6.3.3 Consider the action

$$\phi : K \times K \longrightarrow K,$$

$$\phi(a,x) = ax.$$

ϕ defines a left K-module structure on K.

Definition 6.3.4 Let I be a set and $(V_i)_{i\in I}$ be a family of left K-modules.

$$V = \prod_{i\in I} V_i.$$

The action

$$\phi : (K \times V) \longrightarrow V,$$

$$(a, (x_i)_{i\in I}) \longmapsto (a * x_i)_{i\in I}$$

defines a left K-module structure on V.

6.4 Submodules

Definition 6.4.1 Let V be a left K-module, we call **left sub-K-module** of V any subgroup W of $(V,+)$ such that for any $(a,x) \in K \times W, ax \in W$. (resp. right.)

Example 6.4.2 $\{0\}$ and V itself is a left sub-K-modules of V.

Definition 6.4.3 Let E and F be left-K-modules. We call **homomorphism of left K-modules from E to F** any mapping $f : E \to F$, such that
(1) f is a homomorphism of groups from $(E,+)$ to $(F,+)$.
(2) For any $(a,x) \in K \times E, f(ax) = af(x)$.
If K is communitative, a homomorphism of K-module is also called a K**-linear mapping**.

Lemma 6.4.4 Let V be a left K-module.
(1) $\forall a \in K, a0_V = 0_V$.
(2) $\forall x \in V, 0x = 0_V$.

Proof
(1) $a0_V = a(0_V + 0_V) = a0_V + a0_V \Rightarrow 0_V = a0_V$.
(2) $0x = (0+0)x = 0x + 0x \Rightarrow 0x = 0_V$. □

Theorem 6.4.5 Let $f : E \to F$ be a homomorphism of left-K-modules.
(1) $\ker(f)$ is a left sub-K-module of E.
(2) $\mathrm{Im}(f)$ is a left sub-K-module of F.

Proof First, $\ker(f)$ is a subgroup of E, $\mathrm{Im}(f)$ is a subgroup of F.
(1) Let $a \in K, x \in \ker(f), f(ax) = af(x) = a0_V = 0_V$. So $ax \in \ker(f)$.
(2) Let $y \in \mathrm{Im}(f)$, there exists $x \in E$ such that $f(x) = y$. For any $a \in K, ay = af(x) = f(ax) \in \mathrm{Im}(f)$ □

Proposition 6.4.6 Let V be a left K-module. For any $x \in V, -x = (-1)x$.

Proof

$$(-1)x + x = (-1+1)x = 0x = 0_V.$$

□

Example 6.4.7 Let $(V_i)_{i\in I}$ be a family of left K-modules. We denote by

$$\bigoplus_{i\in I} V_i \text{ the set of } (x_i)_{i\in I} \in \prod_{i\in I} V_i,$$

such that $\{i \in I \mid x_i \neq 0_{V_i}\}$ is finite. This is a subgroup of $\prod_{i\in I} V_i$. For any $a \in K$, and $(x_i)_{i\in I} \in \bigoplus_{i\in I} V_i$,

$$\{i \in I \mid ax_i \neq 0_{V_i}\} \subseteq \{i \in I \mid x_i \neq 0_{V_i}\}.$$

So $a(x_i)_{i\in I} = (ax_i)_{i\in I} \in \bigoplus\limits_{i\in i} V_i$, which means that $\bigoplus\limits_{i\in I} V_i$ is a left sub-K-module of $\prod\limits_{i\in I} V_i$. $\bigoplus\limits_{i\in I} V_i$ is called the direct sum of $(V_i)_{i\in I}$. We denote by

$$K^{\oplus I}$$

the left sub-K-module of K^I.

Proposition 6.4.8 Let E and F be left K-modules, $f : E \to F$ be a mapping.
(1) If f is a homomorphism of left K-modules, for any $n \in \mathbb{N}_{\geq 1}$, any $(a_1, a_2, \ldots, a_n) \in K$, and $(x_1, x_2, \ldots, x_n) \in E^n$,

$$f(a_1x_1 + \cdots + a_nx_n) = a_1f(x_1) + \cdots + a_nf(x_n).$$

(2) Suppose that for any $a \in K$, $(x, y) \in E^2$,

$$f(x + ay) = f(x) + af(y).$$

Then f is a homomorphism of left K-modules.

Proof (1) Induction on n.
(2) Take $a = 1$, for any $(x, y) \in E$, $f(x + y) = f(x) + f(y)$.
Take $x = 0_E$, $f(ay) = 0_F + af(y) = af(y)$. □

Definition 6.4.9 If a left K-module homomorphism is a bijection we say that it is a **left K-module isomorphism**.

6.5 Universal Property

Proposition 6.5.1 Let $(V, +)$ be a communitative group. Then

$$\begin{aligned}\mathbb{Z} \times V &\longrightarrow V\\ (n, x) &\longmapsto nx\end{aligned}$$

defines a $\mathbb{Z}$-module substructure on V.

Proof First, nx is the image of n by the unique homomorphism of groups $\phi_x : \mathbb{Z} \to V, 1 \mapsto x$.

$$(n+m)x = \phi_x(n+m) = \phi_x(n) + \phi_x(m) = nx + mx.$$

Let $(x, y) \in V^2$,

$$\begin{aligned}\phi_x + \phi_y :&\mathbb{Z} \longrightarrow V,\\ &n \longmapsto \phi_x(n) + \phi_y(n) = nx + ny\end{aligned}$$

is a homomorphism of groups, since for any $(n, m) \in \mathbb{Z}^2$

$$\begin{aligned}&(\phi_x + \phi_y)(n+m)\\ =&\phi_x(n+m) + \phi_y(n+m) = \phi_x(n) + \phi_x(m) + \phi_y(n) + \phi_y(m)\\ =&(\phi_x(n) + \phi_y(n)) + (\phi_x(m) + \phi_y(m)).\end{aligned}$$

Since $(\phi_x+\phi_y)(1) = x+y = \phi_{x+y}$, $\phi_{x+y} = \phi_x+\phi_y$. So $n(x+y) = nx+ny, \forall n \in \mathbb{Z}$. $1x = \phi_x(1) = x$. If $n \in \mathbb{N}$,

$$(nm)x = \phi_x(nm) = \phi_x(\underbrace{m + \cdots + m}_{n \text{ copies}}) = n\phi_x(m) = n(mx).$$

If $-n \in \mathbb{N}$,

$$\phi_x(nm) = -\phi_x((-n)m) = -(-n)\phi(m) = n\phi_x(m).$$

□

Proposition 6.5.2 Let V be a left K-module, $x \in V$. There exists a unique homomorphism of left K-modules $\phi_x : K \longrightarrow V$, such that $\phi_x(1) = x$.

Proof If ϕ_x exists, then it should satisfy

$$\forall a \in K, \phi_x(a) = a\phi_x(1) = ax.$$

It suffices to check that $\phi_x : K \to V, a \mapsto ax$ is a homomorphism.

$$\phi_x(a+b) = (a+b)x = ax + bx = \phi_x(a) + \phi_x(b),$$

$$\phi_x(\lambda a) = (\lambda a)x = \lambda(ax) = \lambda\phi_x(a).$$

□

Proposition 6.5.3 Let $(V_i)_{i\in I}$ be a family of left K-modules.
(1) Let W be a left K-module. For any $i \in I$, let $f_i : W \to V_i$ be a homomorphism. Then there exists a unique homomorphism

$$f : W \longrightarrow \prod_{i\in I} V_i,$$

such that

$$\forall i \in I, \pi_i \circ f = f_i,$$

where π_i sends $(x_j)_{j\in I} \in \prod_{j\in I} V_j$ to x_i.
(2) Let W be a left K-module, for any $i \in I$, let $g_i : V_i \to W$ be a homomorphism of left K-modules. There exists a unique homomorphism

$$g : \bigoplus_{i\in I} V_i \longrightarrow W$$

such that

$$\forall i \in I, g \circ \lambda_i = g_i,$$

where

$$\lambda_j : V_j \longrightarrow \bigoplus_{i\in I} V_i,$$

$$x_j \longrightarrow (y_i)_{i\in I} \text{ with } y_i = \begin{cases} x_j, i = j \\ 0, i \neq j \end{cases}.$$

Proof
(1) There exists a unique mapping $f : W \to \prod_{i\in I} V_i$, such that

$$\forall i \in I, \pi_i \circ f = f_i,$$

$$\forall z \in W, f(z) = (f_i(z))_{i\in I}.$$

We have proved that f is a homomorphism of groups.

$$\forall a \in K, z \in W. f(az) = (f_i(az))_{i\in I} = (af_i(z))_{i\in I} = af(z).$$

(2) We have prove that there exists a unique $g : \bigoplus_{i\in I} V_i \to W$ homomorphism of group such that $\forall i \in I, g \circ \lambda_i = g_i$. $g\left((x_i)_{i\in I}\right) = \sum_{i\in I} g_i(x_i)$.

$$\begin{aligned}\forall a \in K, g\left(a(x_i)_{i\in I}\right) =& g\left((ax_i)_{i\in I}\right)\\ &= \sum_{i\in I} g_i(ax_i) = \sum_{i\in I} ag_i(x_i) = a\sum_{i\in I} g_i(x_i) = ag(x).\end{aligned}$$

□

Application 6.5.4 Let V be a left K-module. Let I be a set and $(x_i)_{i\in I} \in V^I$. For any $i \in I$, let

$$\phi_{x_i} : K \longrightarrow V, a \mapsto ax_i.$$

So the family $(\phi_{x_i})_{i\in I}$ determines a homomorphism of left K-modules

$$\Phi : K^{\oplus I} \longrightarrow V,$$

$$(a_i)_{i\in I} \longmapsto \sum_{i\in I} \phi_{x_i}(a_i) = \sum_{i\in I} a_i x_i.$$

6.6 Matrices

Definition 6.6.1 Let $n \in \mathbb{N}$. Let V be a **left** K-module. For any $(x_1, \ldots, x_n) \in V^n$, we denote by

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : K^n \longrightarrow V,$$

$$(a_1, \ldots, a_n) \longmapsto a_1x_1 + \cdots + a_nx_n.$$

This is a homomorphism of left K-modules.

Example 6.6.2 Consider the case where $V = K^p$ with $p \in \mathbb{N}$. Each x_i is of the form $(b_{i,1}, \ldots, b_{i,p})$.

$$\text{So} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ becomes } \begin{pmatrix} b_{1,1} & \ldots & b_{1,p} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \ldots & b_{n,p} \end{pmatrix}.$$

Definition 6.6.3 We call n by p matrix with coefficients in K any homomorphism of left K-module from K^n to K^p.

Definition 6.6.4 Let n and p be natural numbers,and V be a left K-module. Let $A : K^n \to K^p$, and $\varphi : K^p \to V$ be homomorphism of left K-modules. We denote by

$$A\varphi : K^n \longrightarrow V$$

be the mapping $\varphi \circ A$.

Proposition 6.6.5 Let E, F and G be left K-modules. Let $\varphi : E \to F$ and $\psi : F \to G$ be homomorphism of left K-modules. Then $(\psi \circ \varphi) : E \to G$ is a homomorphism of left K-modules.

Proof Let $(x, y) \in E^2, a \in K$. $(\psi \circ \varphi)(x + ay) = \psi(\varphi(x + ay)) = \psi(\varphi(x) + a\varphi(y)) = \psi(\varphi(x)) + a\psi(\varphi(y))$. $\square$

Computation Suppose that

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix}, \ \varphi = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}.$$

For $t = (t_1, \ldots, t_n) \in K^n$,

$$t \overset{A}{\longmapsto} \left(\sum_{i=1}^{n} t_i a_{i,1}, \ldots, \sum_{i=1}^{n} t_i a_{i,p} \right) \overset{\varphi}{\longmapsto} \sum_{j=1}^{p} \sum_{i=1}^{n} t_i a_{i,j} x_j.$$

So,

$$A\varphi = \begin{pmatrix} a_{1,1}x_1 + \cdots + a_{1,p}x_p \\ \vdots \\ a_{n,1}x_1 + \cdots + a_{n,p}x_p. \end{pmatrix}$$

Question Let

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix}, B = \begin{pmatrix} b_{1,1} & \cdots & b_{1,q} \\ \vdots & \ddots & \vdots \\ b_{p,1} & \cdots & b_{p,q} \end{pmatrix}. AB =?$$

We have

$$AB = \begin{pmatrix} a_{1,1}b_{1,1} + \cdots + a_{1,p}b_{p,1} & \cdots & a_{1,1}b_{1,q} + \cdots + a_{1,p}b_{p,q} \\ \vdots & \ddots & \vdots \\ a_{n,1}b_{1,1} + \cdots + a_{n,p}b_{p,1} & \cdots & a_{n,1}b_{1,q} + \cdots + a_{n,p}b_{p,q} \end{pmatrix}.$$

Example 6.6.6 Let $(a_1, a_2, \ldots, a_n) \in K^n$, we denote by

$$\begin{aligned} \operatorname{diag}(a_1, \ldots, a_n) : K^n &\longrightarrow K^n \\ (t_1, \ldots, t_n) &\longmapsto (t_1 a_1, \ldots, t_n a_n). \end{aligned}$$

$\operatorname{diag}(a_1, \ldots, a_n)$ is called a **diagonal matrix**.

Example 6.6.7 $\mathrm{Id}_{K^n} : K^n \longrightarrow K^n,\ t \mapsto t$ is also written as $I_n = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$

Let V be a left K-module, $(x_1, \ldots, x_n) \in V^n$, $(a_1, \ldots, a_n) \in K^n$.

$$\operatorname{diag}(a_1, \ldots, a_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_1 x_1 \\ \vdots \\ a_n x_n \end{pmatrix}.$$

$$\operatorname{diag}(a_1, \ldots, a_n)\operatorname{diag}(b_1, \ldots, b_n) = \operatorname{diag}(a_1 b_1, \ldots, a_n b_n).$$

6.7 Linear Equations

We fix a unitary ring.

Definition 6.7.1 Let $p \in \mathbb{N}$. For $(a_1, \ldots, a_p) \in K^p$, let $j(a_1, \ldots, a_p)$ be the least index $i \in \{1, \ldots, p\}$ such that $a_i \neq 0$. By convention,

$$j(0, \ldots, 0) = p + 1.$$

Let V be a left K-module, $A \in M_{n,p}(K)$. Let $(b_1, \ldots, b_n) \in V^n$. We consider

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \tag{$*$}$$

We write A into the form

$$\begin{pmatrix} \vec{a}^{(1)} \\ \vdots \\ \vec{a}^{(n)} \end{pmatrix}, \vec{a}^{(i)} = (a_{i,1}, \ldots, a_{i,p}).$$

Definition 6.7.2 We say that the matrix is of row echelon form if

$$j\left(\vec{a}^{(1)}\right) \leq j\left(\vec{a}^{(2)}\right) \leq \cdots \leq j\left(\vec{a}^{(n)}\right),$$

and the strict inequality holds once

$$j\left(\vec{a}^{(i)}\right) \leq p.$$

If in addition $a_{i,j\left(\vec{a}^{(i)}\right)} = 1$, and $a_{k,j\left(\vec{a}^{(i)}\right)} = 0$ for any $k \neq i$ once $\vec{a}^{(i)} \neq (0, \ldots, 0)$. We say that A is of **reduced row echelon form**.

Example 6.7.3

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

are of row echelon form.

Theorem 6.7.4 Suppose that A is of reduced echelon form. Let

$$I(A) = \{i \in \{1, \ldots, n\} \mid \vec{a}^{(i)} \neq (0, \ldots, 0)\},$$

$$J_0(A) = \{1, \ldots, p\} \backslash \{j\left(\vec{a}^{(i)}\right) \mid i \in I(A)\}.$$

(1) If there exists $i \in \{1, \ldots, n\} \backslash I(A), b_i \neq 0$ the equation $A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ has no solution.

(2) If $\forall i \in \{1, \ldots, n\} \backslash I(A), b_i = 0$. The solution set of the equation is the image of the following mapping:

$$\Phi : V^{I(A)} \longrightarrow V^p \text{ with}$$

$$(z_l)_{l \in J_0(A)} \longmapsto (x_1, \ldots, x_p),$$

$$x_k = \begin{cases} z_k & \text{, if } k \in J_0(A) \\ b_i - \sum_{l \in J_0(A)} a_{i,l} z_l & \text{, if } k = j\left(\vec{a}^{(i)}\right). \end{cases}$$

Proposition 6.7.5 Let m, n, p be natural numbers. $S \in M_{m,n}(K)$, $A \in M_{n,p}$. If $(x_1, \ldots, x_p)$ is a solution of the equation

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}. \tag{$*$}$$

Then it is also a solution of the equation

$$(SA) \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = S \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}. \tag{$*_S$}$$

Moreover, if S is left invertible (namely there exists $T \in M_{n,m}(K)$ such that $TS = I_n$), then $(*)$ and $(*_S)$ have the same solution set.

Proof

$$(SA) \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = S \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

So

$$TSA \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = TS \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \Rightarrow A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

□

Definition 6.7.6 Let $n \in \mathbb{N}$ and $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ be a bijection.

Denote by

$$P_\sigma : K^n \longrightarrow K^n,$$

$$P(t_1, \dots, t_n) := \left(t_{\sigma^{-1}(1)}, \dots, t_{\sigma^{-1}(n)}\right).$$

P_σ is a homomorphism of left K-modules.

$$P_\sigma P_{\sigma^{-1}} = P_{\sigma^{-1}} P_\sigma = I_n.$$

Let V be a left K-module, $(x_1, \dots, x_n) \in V$,

$$P_\sigma \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : K^n \longrightarrow V,$$

$$(t_1, \dots, t_n) \xmapsto{P_\sigma} (t_{\sigma^{-1}(1)}, \dots, t_{\sigma^{-1}(n)}) \xmapsto{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}} \sum_{i=1}^{n} t_{\sigma^{-1}(i)} x_i = \sum_{j=1}^{n} t_j x_{\sigma(j)}.$$

$$P_\sigma \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_{\sigma(1)} \\ \vdots \\ x_{\sigma(n)} \end{pmatrix}, \; P_\sigma \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$$

Definition 6.7.7 If $\underline{r} = (r_1, r_2, \dots, r_n) \in K^n$, we denote by $D_{\underline{r}}$ the matrix $\operatorname{diag}(r_1, \dots, r_n)$. If for any $i \in \{1, \dots, n\}$, r_i is left invertible and is a inverse of s_i, then

$$D_{\underline{s}} D_{\underline{r}} = I_n.$$

Definition 6.7.8 Let $n \in \mathbb{N}, i \in \{1, \dots, n\}, c = (c_1, \dots, c_n) \in K^n, c_i = 0$. Denote by

$$S_{i,c} : K^n \longrightarrow K^n,$$

$$S_{i,c}(t_1, \dots, t_n) := \left(t_1, \dots, t_{i-1}, t_i + \sum_{j=1}^{n} t_j c_j, t_{i+1}, \dots, t_n\right)$$

$$S_{i,c}S_{i,-c} = S_{i,-c}S_{i,c} = I_n$$

$$S_{i,c}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : (t_1, \dots, t_n) \longmapsto \sum_{j=1}^{n} t_j x_j + \sum_{j=1}^{n} t_j c_j x_i$$

$$S_{i,c}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 + c_1 x_i \\ \vdots \\ x_i \\ \vdots \\ x_n + c_n x_i \end{pmatrix}$$

Definition 6.7.9 Let $G_n(K)$ be the subset of $M_{n,n}(K)$ consisting of matrices S, that can be written as $U_1, \dots, U_N$, where $N \in \mathbb{N}$ (if $N = 0$, by convention, $S = I_n$) and each U_i is of the following forms:
(1) P_σ, with $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ being a bijection.
(2) $D_{\underline{r}}$ with each r_i being left invertible.
(3) $S_{i,c}$ with $i \in \{1, \dots, n\}, c = (c_1, \dots, c_n) \in K^n, c_i = 0$.
Let $p \in \mathbb{N}$. We say that $A \in M_{n,p}(K)$ is **reducible by Gaussian elimination** if there exists $S \in G_n(K)$ such that SA is of reduced row echelon form.

Lemma 6.7.10 If $A \in M_{n,p}(K)$ is such that SA is reducible by Gaussian elimination, for some $S \in G_n(K)$, then A is also reducible by Gaussian elimination.

Theorem 6.7.11 Suppose that K is a division ring. For any $(n,p) \in \mathbb{N}^2$, any matrix $A \in M_{n,p}(K)$ is reducible by Gaussian elimination.

Proof We reason by induction on p.
$p = 0$. A is already of reduced row echelon form.
Suppose that the statement is true for matrices of at most $p-1$ columns. ($p \geq 1$)
We write A as $\begin{pmatrix} \lambda_1 & \\ \vdots & B \\ \lambda_n & \end{pmatrix}$ where $B \in M_{n,p-1}(K)$. If $\lambda_1 = \cdots = \lambda_n = 0$. By induction hypothesis, there exists $S \in G_n(K)$ such that SB is of reduced row

echelon form.

$$SA = \begin{pmatrix} 0 & & \\ \vdots & SB & \\ 0 & & \end{pmatrix}$$

is of reduced row echelon form. If $(\lambda_1, \ldots, \lambda_n) \neq (0, \ldots, 0)$, by the lemma, we may suppose that $\lambda_1 \neq 0$ (By permuting rows). By multiplying A by $\mathrm{diag}(\lambda_1^{-1}, 1, \ldots, 1)$ we may assume (by the lemma) that $\lambda_1 = 1$. So

$$A = \begin{pmatrix} 1 & & \\ \lambda_2 & & \\ & B & \\ \vdots & & \\ \lambda_n & & \end{pmatrix}.$$

By multiplying $S_{1,(0,-\lambda_2,\ldots,-\lambda_n)}$ and A, we may assume (by the lemma) that A is of the form

$$\begin{pmatrix} 1 & \mu_2 & \ldots & \mu_n \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix}.$$

Applying the induction hypothesis to C. (For any $T \in G_{n-1}(K), T : K^{n-1} \to K^{n-1}, S : K^n \to K^n$, $S(t_1, \ldots, t_n) = (t_1, T(t_2, \ldots, t_n))$ belongs to $G_n(K)$.) We write C as $\begin{pmatrix} c_2 \\ \vdots \\ c_n \end{pmatrix}$ where $c_2, \ldots, c_k$ belong to $k^{p-1} \backslash \{(0, \ldots, 0)\}, c_{k+1} = \cdots = c_n = (0, \ldots, 0), j(c_2) < \cdots < j(c_k)$. For any $i \in \{2, \ldots, k\}$, we multiply $-\mu_{j(c_i)}$ times the i^{th} row of A to the first row. The result is a matrix of reduced row echelon form. $\square$

6.8 Quotient Modules

Let K be a unitary ring.

Proposition 6.8.1 Let E be a left K-module, F be a left sub-K-module of E. The mapping

$$K \times E/F \longrightarrow E/F,$$

$$(a, [x]) \longmapsto [ax]$$

(Resp. right,$[xa]$) is well defined, and determines a structure of left K-module on

E/F. Moreover, the projection mapping

$$\pi : E \longrightarrow E/F$$

$$x \longmapsto [x]$$

is a homomorphism.

Proof Recall that F is a subgroup of $(E, +)$ such that

$$\forall a \in K, \forall y \in F, ay \in F,$$

$$[x] = \{y \in E \mid y - x \in F\}.$$

If $[x] = [y]$, then $y - x \in F$, so $ay - ax = a(y - x) \in F$, which means $[ay] = [ax]$.
(1) $[1x] = [x]$.
(2) $(ab)[x] = [(ab)x] = [a(bx)] = a[bx] = a(b[x])$.
(3)

$$(a + b)[x] = [(a + b)x] = [ax + bx] = [ax] + [bx] = a[x] + b[x].$$

$$a[x + y] = [a(x + y)] = [ax + ay] = [ax] + [ay] = a[x] + a[y].$$

Finally,

$$\pi(x + ay) = [x + ay] = [x] + [ay] = [x] + a[y] = \pi(x) + a\pi(y).$$

□

Theorem 6.8.2 Let $f : V \to W$ be a homomorphism of left K-modules.
(1) $\mathrm{Im}(f)$ is a sub-K-module of W.
(2) $\ker(f)$ is a sub-K-module of V.
(3) $\tilde{f} : V/\ker(f) \longrightarrow W, [x] \longmapsto f(x)$ is a homomorphism of left K-modules. Moreover, as a mapping, $\tilde{f}$ is injective and has $\mathrm{Im}(f)$ as its range. Hence it defines an isomorphism between $V/\ker(f)$ and $\mathrm{Im}(f)$.

Proof
(1) We have proved that $\mathrm{Im}(f)$ is a subgroup of W. If $y = f(x) \in \mathrm{Im}(f), \forall a \in K, ay = af(x) = f(ax) \in \mathrm{Im}(f)$. So $\mathrm{Im}(f)$ is a left sub-K-module of W.
(2) We have proved that $\ker(f)$ is a subgroup of V. If $x \in \ker(f), \forall a \in K, f(ax) = af(x) = a0 = 0$. So $\ker(f)$ is a left sub-K-module of V.

(3) We have proved that $\tilde{f}$ is an injective homomorphism of groups, with $\mathrm{Im}(\tilde{f}) = \mathrm{Im}(f)$. So $\tilde{f}$ defines an isomorphism of group $V/\ker(f) \longrightarrow \mathrm{Im}(f)$. Moreover, $\tilde{f}(a[x]) = \tilde{f}([ax]) = f(ax) = af(x) = a\tilde{f}([x])$. So $\tilde{f}$ is a homomorphism of left K-modules. $\square$

6.9 Quotient Ring

Proposition 6.9.1 Let A be a unitary ring. Let $\sim$ be an equivalence relation on A that is compatible with the addition and with the multiplication. Then $A/\sim$ equipped with the quotient composition law of $+$ and $\cdot$ forms a unitary ring, and the projection mapping $\pi : A \longrightarrow A/\sim$ is a homomorphism of unitary ring.

Proof We have seen that $(A/\sim, +)$ forms an abelian group and $(A, \cdot)$ forms a monoid, and $\pi : A \longrightarrow A/\sim$ is a homomorphism of additive groups and multiplicative monoids. It remains to check the distributivity.

$$[a]([b] + [c]) = [a][b + c] = [a(b + c)] = [ab + ac] = [ab] + [ac] = [a][b] + [a][c].$$

$$([b] + [c])[a] = [(b + c)a] = [ba + ca] = [b][a] + [c][a].$$

$\square$

Definition 6.9.2 $A/\sim$ is called the **quotient ring of** A .

Remark 6.9.3 There exists a subgroup I of A such that

$$a \sim b \Leftrightarrow b - a \in I.$$

$\forall x \in I, [x] = 0$, so for any $a \in A$,

$$[ax] = [a][x] = 0, \ [xa] = [x][a] = 0.$$

So I is a left sub-A-module of A and a right sub-A-module of A.

Definition 6.9.4 Let A be a unitary ring. If a subset I of A is a left sub-A-module of A and a right sub-A-module of A, then we call I a **ideal** of A. If I is an ideal of A, then the composition laws of A define by passing to quotient a structure of unitary ring on the quotient mapping A/I. So that A/I becomes a quotient ring of A.

Theorem 6.9.5 Let $f : A \to B$ be a homomorphism of unitary rings. Let $I = \ker(f)$.
(1) I is an ideal of A.
(2) $f(A)$ is a unitary subring of B.
(3) f induces $\tilde{f} : A/I \longrightarrow f(A)$ an isomorphism of unitary rings.

Proof
(1)

$$\forall a \in A, \forall x \in I, f(ax) = f(a)f(x) = f(a)0 = 0 = 0f(a) = f(x)f(a) = f(xa).$$

So $\{ax, xa\} \subseteq I$. Since I is a subgroup of A, it is actually an ideal.
(2) Since f is a homomorphism of groups $(A, +) \longrightarrow (B, +)$ and a homomorphism of monoids $(A, \cdot) \longrightarrow (B, \cdot)$, $f(A)$ is a subgroup of $(B, +)$ and a submonoid of $(B, \cdot)$.
(3) $\tilde{f}$ is a homomorphism of unitary rings. $\tilde{f}([x]) := f(x)$. In the same time $\tilde{f} : A/I \longrightarrow f(A)$ is a bijection. So it is a homomorphism of rings. □

Example 6.9.6 Consider $\mathbb{Z}$. Let I be an ideal of $\mathbb{Z}$. If $I \neq \{0\}$, then $I \cap \mathbb{N}_{\geq 1} \neq \varnothing$. Let $d \in I \cap \mathbb{N}_{\geq 1}$ be the least element. For any $n \in I$, we can write n as

$$n = dm + r, \text{ where } m \in \mathbb{Z}, r \in \{0, \ldots, d-1\}.$$

So $r = n - dm \in I$, which means $r = 0$. Therefore, $I = d\mathbb{Z}$.

Definition 6.9.7 Let A be a communitative unitary ring. If an ideal of A is of the form

$$Ax : \{ax \mid a \in A\} \text{ with } x \in A.$$

We say that it is a **principal ideal**. If all ideals of A are principal, we say that A is a **principal ideal ring**.

Example 6.9.8 $\mathbb{Z}$ is a principal ideal ring.

Remark 6.9.9 If A is a unitary ring, $\mathbb{Z} \longrightarrow A$, $n \longmapsto n1_A$ is the unique homomorphism of unitary rings. $\ker(\mathbb{Z} \longrightarrow A)$ is an ideal of $\mathbb{Z}$. It is of the form

$d\mathbb{Z}, d \in \mathbb{N}$. This natural number d is called the **characteristic** of A, denoted as $\operatorname{char}(A)$.

Definition 6.9.10 Let A be a communitative unitary ring. Let $a \in A$. If $\exists b \in A \backslash\{0\}$ such that $ab = 0$, we say that a is a zero divisor. If $0 \in A$ is the ONLY zero divisor, we say that A is an **(integral) domain**.

A is an integral domain if and only if $0 \neq 1$, and $\forall (a,b) \in (A\backslash\{0\})^2, \ ab \neq 0$.

Example 6.9.11
$\mathbb{Z}$ is an integral domain.
All field are integral domains.
$\mathbb{Z}/6\mathbb{Z}$ is NOT a integral domain.$[2][3] = [6] = [0]$.

Proposition 6.9.12 All unitary subrings of an integral domain are integral domains.

Proposition 6.9.13 Let A be a unitary ring. E be a left A-module and I be an ideal of A. Suppose that

$$\forall (a,x) \in I \times E, \ ax = 0. \ (I \text{ annihilates } E)$$

Then the mapping

$$(A/I) \times E \longrightarrow E,$$
$$([a], x) \longmapsto ax$$

is well defined and defines a left A-module structure on E.

Proof If $[a] = [b]$, then $b - a \in I$. So $\forall x \in E, (b-a)x = bx = ax = 0$. Hence $ax = bx$. $\forall (a,b) \in A \times A, \forall (x,y) \in E \times E$:
(1) $[1]x = 1x = x$. $([a][b])\,x = [ab]x = (ab)x = a(bx) = [a](bx) = [a]([b]x)$.
(2) $([a]+[b])\,x = [a+b]x = (a+b)x = ax+bx = [a]x+[b]x$. $[a](x+y) = a(x+y) = ax+ay = [a]x+[a]y$. □

6.10 Free Modules

We fix a unitary ring K.

Definition 6.10.1 Let V be a left K-module. For any family $\underline{x} := (x_i)_{i \in I} \in V^I$, we denote by

$$\varphi_{\underline{x}} : K^{\oplus I} \longrightarrow V$$

the homomorphism sending $(a_i)_{i \in I}$ to $\sum_{i \in I} a_i x_i$.
(1) $\mathrm{Im}\,(\varphi_{\underline{x}})$ is a left K-submodule of V, called the **left sub-K-module generated by** $\underline{x}$, denote as $\mathrm{Span}_K\,((x_i)_{i \in I})$. If φ_x is surjective, we say that $(x_i)_{i \in I}$ is a **system of generators** of V. $(\forall y \in V, \exists (a_i)_{i \in I} \in K^{\oplus I}, y = \sum_{i \in I} a_i x_i)$ Elements of $\mathrm{Span}_K\,((x_i)_{i \in I})$ are called **K-linear combinations** of $(x_i)_{i \in I}$.
(2) If φ_x is injective, we say that $(x_i)_{i \in I}$ is **K-linearly independent**. $(\forall (a_i)_{i \in I} \in K^{\oplus I}, \sum_{i \in I} a_i x_i = 0 \to a_i = 0, \forall i \in I)$
(3) If $\varphi_{\underline{x}}$ is an isomorphism, we say $(x_i)_{i \in I}$ is a **basis** of V. If V has at least a basis, we say that V is a **free left K-module**. If V has a system of generators $(x_i)_{i \in I}$ such that I is finite, we say that V is **finitely generated**, or is **finite types**.

Example 6.10.2 $K^{\oplus I}$ is a free left K-module.

Remark 6.10.3 Any left K-module is isomorphic to a free quotient module of a free left K-module.

Theorem 6.10.4 Let K be a division ring and V be a left K-module of finite type. Let $(x_i)_{i=1}^n$ be a system of generators of V. There exists $I \subseteq \{1, \ldots, n\}$ such that $(x_i)_{i \in I}$ forms a basis of V.

Proof By induction on n.
Case $n = 0$, $V = \{0\}.(x_i)_{i \notin \varnothing}$ is a basis of V. Suppose that $n \geq 1$. If $(x_i)_{i=1}^n$ is K-linearly independent, it is already a basis. Otherwise there exists $0 \neq (b_1, \ldots, b_n) \in K^n$ such that $b_1 x_1 + \cdots + b_n x_n = 0$. By permuting $x_1, \ldots, x_n$, we may assume that $b_n \neq 0$. $x_n = -b_n^{-1}(b_1 x_1 + \cdots + b_{n-1} x_{n-1})$. For any $y \in V$, there exists $(a_1, \ldots, a_n) \in K^n$, such that

$$y = \sum_{i=1}^{n} a_i x_i = \sum_{i=1}^{n-1} a_i x_i - a_n b_n^{-1} \left(b_1 x_1 + \cdots + b_{n-1} x_{n-1}\right).$$

□

Theorem 6.10.5 Let K be a unitary ring, V be a left K-module and W be a left sub-K-module of V. Let $(x_i)_{i=1}^n \in W^n$ and $(\alpha_j)_{j=1}^l \in (V/W)^l$, with $(n, l) \in \mathbb{N}^2$.

For any $j \in \{1, \ldots, l\}$. Let x_{n+j} be an element of the equivalence class of α_j. ($[x_{n+j}] = \alpha_j$)
(1) If $(x_i)_{i=1}^n$ and $(\alpha_j)_{j=1}^l$ are K-linearly independent, then $(x_i)_{i=1}^{n+l}$ is K-linearly independent.
(2) If $(x_i)_{i=1}^n$ and $(\alpha_j)_{j=1}^l$ are system of generators, then $(x_i)_{i=1}^{n+l}$ is a system of generators.

Proof
(1) Let $(a_i)_{i=1}^{n+l} \in K^{n+l}$ such that

$$\sum_{i=1}^{n+l} a_i x_i = 0.$$

Taking the equivalence class of both sides in V/W, we get $\sum_{j=1}^{l} a_{n+j}\alpha_j = [0]$. So $a_{n+1} = \cdots = a_{n+l} = 0$. Hence $a_1 x_1 + \cdots + a_n x_n = 0$. So $a_1 = \cdots = a_n = 0$.
(2)Let $y \in V$. There exists $(c_{n+1}, \ldots, c_{n+l}) \in K^l$, such that

$$[y] = c_{n+1}\alpha_1 + \cdots + c_{n+l}\alpha_l = [c_{n+1}x_{n+1} + \cdots + c_{n+l}x_{n+l}].$$

So, $y-(c_{n+1}x_{n+1} + \cdots + c_{n+l}x_{n+l}) \in W$. Hence, there exists $(c_1, \ldots, c_n) \in K^n$,

$$y - (c_{n+1}x_{n+1} + \cdots + c_{n+l}x_{n+l}) = c_1 x_1 + \cdots + c_n x_n.$$

So $y = \sum_{i=1}^{n+l} c_i x_i$. $\square$

Proposition 6.10.6
(1) If A is injective and $(x_i)_{i\in I}$ is a K-linearly independent, then $(y_j)_{j\in I}$ is K-linearly independent.
(2) If A is surjective, then $(x_i)_{i\in I}, (y_j)_{j\in J}$ generate the same left sub-K-module of V.

$$\mathrm{Im}(\varphi_y) = \mathrm{Im}(\varphi_{\underline{x}} \circ A) = \mathrm{Im}(\varphi_{\underline{x}}).$$

In particular, if $f(x_i)_{i\in I}$ is a system of generators, and A is surjective, then $(y_j)_{j\in J}$ is a system of generators.
(3) If $(x_i)_{i\in I}$ is a basis and A is a bijection, then $(y_j)_{j\in J}$ is a basis.

Application Let $n \in \mathbb{N}$, $(x_1, \ldots, x_n) \in V^n$. Let $(y_1, \ldots, y_n) \in V^n$ such that

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = S \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ with } S \text{ invertible.}$$

$(x_i)_{i\in I}^n$ is K-linearly independent if and only if $(y_j)_{j\in J}^n$ is K-linearly independent.
$(x_i)_{i\in I}^n$ is a system of generators if and only if $(y_j)_{j\in J}^n$ is a system of generators.
$(x_i)_{i\in I}^n$ is a basis if and only if $(y_j)_{j\in J}^n$ is a basis.

Theorem 6.10.7 Let $(n,p) \in \mathbb{N}^2$ and $A \in M_{n,p}(K)$. We write A into the form

$$A = \begin{pmatrix} \underline{a}^{(1)} \\ \ldots \\ \underline{a}^{(n)} \end{pmatrix} \text{ where } \underline{a}^{(i)} = (a_{i,1}, \ldots, a_{i,p}) \in K^p.$$

Assume that A is of reduced row echelon form.
(1) $\left(\underline{a}^{(i)}\right)_{i=1}^n$ is K-linearly independent if and only if $\forall i \in \{1, \ldots, n\}, \underline{a}^{(i)} \neq (0, \ldots, 0)$.
(2) $\left(\underline{a}^{(i)}\right)_{i=1}^n$ is a system of generators if and only if there are exactly p non-zero elements among $\underline{a}^{(1)}, \ldots, \underline{a}^{(n)}$.

Proof
(1) It suffice to check, if $\forall i \in \{1, 2, \ldots, n\}, \underline{a}^{(i)} \neq (0, \ldots, 0)$, then $\left(\underline{a}^{(i)}\right)_{i=1}^n$ is K-linearly independent. Since A is of reduced row echelon form

$$1 \leq j(\underline{a}^{(1)}) < j(\underline{a}^{(2)}) < \cdots < j(\underline{a}^{(n)}) \leq p.$$

Suppose that $(\lambda_1, \ldots, \lambda_n) \in K^n$ such that

$$\lambda_1 \underline{a}^{(1)} + \cdots + \lambda_n \underline{a}^{(n)} = (0, \ldots, 0).$$

Note that the coordinate of index $j(\underline{a}^{(i)})$ of $\lambda_1 \underline{a}^{(1)} + \cdots + \lambda_n \underline{a}^{(n)}$ is λ_i, so

$$\lambda_1 = \cdots = \lambda_n = 0.$$

(2) "$\Leftarrow$": Suppose that $\underline{a}^{(i)} \neq (0, \ldots, 0)$ for $i \in \{1, \ldots, p\}$. Since $1 \leq j(\underline{a}^{(1)}) < \cdots < j(\underline{a}^{(p)}) \leq p$, one has $j(\underline{a}^{(i)}) = i, \forall i \in \{1, \ldots, p\}$. Hence

$$\lambda_1 \underline{a}^{(1)} + \cdots + \lambda_p \underline{a}^{(p)} = (\lambda_1, \ldots, \lambda_p).$$

"⇒": suppose that $(a_i)_{i=1}^n$ is a system of generators. There could not be more than p non-zero elements among $\underline{a}^{(1)}, \ldots, \underline{a}^{(n)}$. If $\underline{a}^{(1)}, \ldots, \underline{a}^{(k)}$ are non-zero and

$$\underline{a}^{(k+1)} = \cdots = \underline{a}^{(n)} = 0,$$

let $(b_1, \ldots, b_p) \in K^p \backslash \{(0, \ldots, 0)\}$, $\forall i \in \{1, \ldots, k\}, b_{j(\underline{a}^{(1)})} = 0$. If $(b_1, \ldots, b_p)$ is a linear combination of $\underline{a}^{(1)}, \ldots, \underline{a}^{(n)}$, there exists $(\lambda_1, \ldots, \lambda_k)$ such that

$$\lambda_1 \underline{a}^{(1)} + \cdots + \lambda_k \underline{a}^{(k)} = (b_1, \ldots, b_p).$$

So $\lambda_1 = \cdots = \lambda_k = 0$. $\square$

Definition 6.10.8 Let K be a division ring and V is a left K-module of finite type. We denote by $\mathrm{rk}_K(V)$ or $\mathrm{rk}(V)$ the least cardinality of the bases V, called the **rank** of V. If K is a field, then $\mathrm{rk}(V)$ is also denoted as $\dim(V)$, called the **dimension** of V. If $f : W \longrightarrow V$ is a homomorphism of left K-modules, the rank of f is defined as the rank of $\mathrm{Im}(f)$, denoted as $\mathrm{rk}(f)$.

Theorem 6.10.9 (rank-nullity theorem) Let K be a division ring and V be a left K-module of finite type, and W be a left sub-K-module of V.
(1) W and V/W are of finite type, and $\mathrm{rk}(W) + \mathrm{rk}(V/W) = \mathrm{rk}(V)$.
(2) Any basis of V has $\mathrm{rk}(V)$ as its cardinality.

Proof
(1) Let $(x_i)_{i=1}^n$ be a basis of V. Then $([x_i])_{i=1}^n$ also form a system of generators of V/W. By theorem 6.10.4, one can extract a subset $I \subseteq \{1, 2 \ldots, n\}$ such that $([x_i]_{i\in I})$ forms a basis of V/W. By permuting the elements $x_1, x_2, \ldots, x_n$, we may assume, without loss of generality, that $I = \{1, 2, \ldots, l\}$, $l \leq n$. For any $j \in \{l+1, \ldots, n\}$ there exists $(b_{j,1}, \ldots, b_{j,l})$ such that

$$[x_j] = \sum_{i=1}^{l} b_{j,i}[x_i].$$

$$y_j := x_j - \sum_{i=1}^{l} b_{j,i} x_i.$$

For any $x \in W$, there exists $(a_i)_{i=1}^n \in K^n$ such that

$$x = \sum_{i=1}^{l} a_i x_i + \sum_{j=l+1}^{n} a_j \left(y_j + \sum_{i=1}^{l} b_{j,i} x_i \right) = \sum_{i=1}^{l} \left(a_i + \sum_{j=l+1}^{n} a_j b_{j,i} \right) x_i + \sum_{j=l+1}^{n} a_j y_j.$$

Taking the equivalence class of $x \in V/W$ (i.e.[0]) we obtain.

$$\forall i \in \{1, \dots, l\},\ a_i + \sum_{j=l+1}^{n} a_j b_{i,j} = 0.$$

Hence,

$$x = \sum_{j=l+1}^{n} a_j y_j.$$

Therefore, W is of finite type, and $\mathrm{rk}(W) + \mathrm{rk}(V/W) \leq \mathrm{rk}(V)$. Moreover, by theorem 6.10.5,

$$\mathrm{rk}(V) \leq \mathrm{rk}(W) + \mathrm{rk}(V/W).$$

Hence,

$$\mathrm{rk}(W) + \mathrm{rk}(V/W) = \mathrm{rk}(V).$$

(2) We reason by induction on $\mathrm{rk}(V)$. If $\mathrm{rk}(V) = 0$, then $\{\varnothing\}$ is the only basis. If $\mathrm{rk}(V) = 1$, then V is of the form Ke, where e is a non-zero element of V. Suppose $\mathrm{rk}(V) = n \geq 1$, and the statement has been proven for modules of $\mathrm{rk} < n$. Let $(e_i)_{i=1}^{m}$ be a basis of V. Let $W = K \cdot e_1$. Then, $([e_i])_{i=2}^{m}$ forms a system of generators of V/W. Moreover, $(a_i)_{i=2}^{m} \in K^{m-1}$ such that

$$\sum_{i=2}^{m} a_i[e_i] = 0,$$

then,

$$\sum_{i=2}^{m} a_i e_i \in W,$$

and hence, there exists $a_1 \in K$,

$$\sum_{i=1}^{m} a_i e_i = 0.$$

We conduct that, in particular

$$a_2 = \cdots = a_n = 0.$$

Hence $([e_i])_{i=2}^{m}$ is a basis of V/W, $\mathrm{rk}(V/W) = n-1$, so $n-1 = m-1 \Leftrightarrow n = m$. $\square$

6.11 Algebra

In this section, we fix a communicative unitary ring K.

Definition 6.11.1 Let K be a communicative unitary ring. If A is a K-module equipped with a composition law

$$A \times A \longrightarrow A,$$

$$(a, b) \longmapsto ab.$$

such that $(A, +, \cdot)$ forms a unitary ring, such that

$$\forall \lambda \in K, \forall (a, b) \in A \times A,\ \lambda (ab) = (\lambda a) b = a (\lambda b).$$

Then we say that A is a **K-Algebra**.

Remark 6.11.2

$$K \longrightarrow A,$$

$$\lambda \longmapsto \lambda 1_A.$$

is a homomorphism of unitary rings.
(1) $(\lambda + \mu)1_A = \lambda 1_A + \mu 1_A$.
(2) $(\lambda 1_A)(\mu 1_A) = \lambda (1_A(\mu 1_A)) = \lambda (\mu(1_A 1_A)) = \lambda(\mu 1_A) = (\lambda\mu)1_A$.
(3) $1_K 1_A = 1_A$.

Remark 6.11.3 Suppose that A is a unitary ring and $f : K \longrightarrow A$ be a homomorphism of unitary rings such that $\forall \lambda \in K,\ \forall a \in A\ af(\lambda) = f(\lambda)a$. Then

$$K \times A \longrightarrow A,$$

$$(\lambda, a) \longmapsto f(\lambda)a$$

defines a structure of K-modules on A.

$$f(\lambda\mu)a = f(\lambda)f(\mu)a = f(\lambda)(f(\mu)a),$$
$$f(1_K)a = 1_A a = a,$$
$$f(\lambda + \mu)a = (f(\lambda) + f(\mu))a = f(\lambda)a + f(\mu)a,$$
$$f(\lambda)(a + b) = f(\lambda)a + f(\lambda)b.$$

Moreover,

$$f(\lambda)(ab) = (f(\lambda)a)b = a(f(\lambda)b).$$

Therefore, A equipped with a structure of K-algebra.

Example 6.11.4 (1) $\{0\}$, (2) K.

Example 6.11.5 Let $(S, \cdot)$ be a monoid. We denote by $K[\![S]\!]$ the K-module K^S. If $(a_s)_{s\in S}$ belongs to K^S, while coordinating $(a_s)_{s\in S}$ as an element of $K[\![S]\!]$, we write it formally as

$$\sum_{s\in S} a_s s.$$

Assume that, for any $s \in S$, the preimage of s by the mapping

$$S \times S \longrightarrow S,$$

$$(\alpha, \beta) \longmapsto \alpha\beta$$

is finite.

$$(\{(\alpha, \beta) \in S \times S \mid \alpha\beta = s\} \text{ is finite.})$$

For example,

$$\mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N},$$

$$(m, n) \longmapsto m + n.$$

$$\forall k \in \mathbb{N}, \{(m, n) \in \mathbb{N} \times \mathbb{N} \mid m + n = k\} \text{ is finite.}$$

We define a composition law on $K[\![S]\!]$ by

$$K[\![S]\!] \times K[\![S]\!] \longrightarrow K[\![S]\!]$$

$$\left(\sum_{s\in S} a_s s, \sum_{s\in S} b_s s\right) \longmapsto \sum_{s\in S}\left(\sum_{(u,v)\in S^2, uv=s} a_u b_v\right) s.$$

We write $(\mathbb{N}, +)$ formally as

$$(\{T^n \mid n \in \mathbb{N}\}, \cdot)$$

such that $T^n \cdot T^m := T^{n+m}$. In this particular case, we write $K[\![N]\!]$ as $K[\![T]\!]$. The element of $K[\![T]\!]$ is of the form

$$\sum_{n\in\mathbb{N}} a_n T^n.$$

It is called a **formal power series (of variable T) with coefficients in K**.

Proposition 6.11.6 $K[\![S]\!]$ is a K-algebra.

Proof

$$\begin{aligned}
&\sum_{s\in S} a_s s\left(\left(\sum_{s\in S} b_s s\right)\left(\sum_{s\in S} c_s s\right)\right)\\
=&\left(\sum_{s\in S} a_s s\right)\left(\sum_{s\in S}\left(\sum_{vw=s} b_v c_w\right) s\right)\\
=&\sum_{s\in S}\left(\sum_{uvw=s} a_u b_v c_w\right) s\\
=&\left(\sum_{s\in S} a_s s\right)\left(\sum_{s\in S} b_s s\right)\left(\sum_{s\in S} c_s s\right)
\end{aligned}$$

□

Definition 6.11.7 Let A be a K-algebra. If B is a subset of A which is a sub-K-module and a unitary subring of A, we say that B is a **sub-K-algebra** of A.

Example 6.11.8 Let S be a monoid. We write $K^{\oplus S}$ as $k[S]$ and define

$$K[S]\times K[S]\longrightarrow K[S],$$

$$\left(\left(\sum_{s\in S} a_s s\right),\left(\sum_{s\in S} b_s s\right)\right)\longmapsto \sum_{s\in S}\left(\sum_{uv=s} a_u b_v\right) s.$$

Then $K[S]$ forms a K-algebra. If $K[\![S]\!]$ is well defined, then $K[S]$ ia a sub-K-algebra of $K[\![S]\!]$.

Proposition 6.11.9 If K is a integral domain, then so is $K[\![T]\!]$.

Proof Let $F=\sum_{n\in\mathbb{N}} a_n T^n$, $G=\sum_{n\in\mathbb{N}} b_n T^n$ be non-zero elements of $K[\![T]\!]$. Let k and l be respectively the least indices such that $a_k\neq 0$, $b_l\neq 0$. We write FG in the form $\sum_{n\in\mathbb{N}} c_n T^n$. Then $c_{k+l}=\sum_{i+j=k+l} a_i b_j = a_k a_l \neq 0$. □

Proposition 6.11.10 Let $F=\sum_{n\in\mathbb{N}} a_n T^n \in K[\![T]\!]$, then $F\in K[\![T]\!]^{\times}$ if and only if $a_0\in K^{\times}$.

Proof "⇒": Suppose $G=\sum_{n\in\mathbb{N}} b_n T^n$ such that $FG=1$. Write FG into the form $\sum_{n\in\mathbb{N}} c_n T^n$. Then $c_0=a_0b_0=1$, so $a_0\in K^{\times}$.

"$\Leftarrow$": Suppose $a_0 \in K^\times$. We want to construct $G = \sum_{n\in\mathbb{N}} b_n T^n$ such that $FG = \sum_{n\in\mathbb{N}} \left(\sum_{i=0}^n a_i b_{n-i}\right) T^n = 1$. One should have $b_0 = a_0^{-1}$, and when $n > 0$,

$$a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_1 + a_n b_0 = 0.$$

Namely, $b_n = -a_0^{-1}(a_1 b_{n-1} + \cdots + a_n b_0)$. We then construct recursively a sequence $(b_n)_{n\in\mathbb{N}}$ by taking $b_0 = a_0^{-1}$ and $b_{n+1} = -a_0^{-1}(a_1 b_n + \cdots + a_n b_0)$. Then $G = \sum_{n\in\mathbb{N}} b_n T^n$ is the inverse of F. □

Example 6.11.11 Let $a \in K$, $(1+aT)^{-1} = \sum_{n\in\mathbb{N}} (-a)^n T^n$.

Definition 6.11.12 Suppose that, for any $n \in \mathbb{N}_{\geq 1}$, $n1_K$ is invertible in K. We denote by $\frac{1}{n}a$ the element $(n1_K)^{-1} a$ in K. We define, for any $a \in K$, an element $\exp(aT) = \sum_{n\in\mathbb{N}} \dfrac{a^n}{n!} T^n$ in $K[\![T]\!]$.

Proposition 6.11.13 For any $(a,b) \in K \times K$,

$$\exp(aT)\exp(bT) = \exp\left((a+b)T\right).$$

So

$$\begin{aligned}(K,+) &\longrightarrow (K[\![T]\!]^\times, \cdot),\\ a &\longmapsto \exp(aT)\end{aligned}$$

is a homomorphism of groups.

Proof

$$\begin{aligned}\exp(aT)\exp(bT) &= \sum_{n\in\mathbb{N}} \left(\sum_{i=0}^n \frac{a^i}{i!}\frac{b^{n-i}}{(n-i)!}\right) T^n \\ &= \sum_{n\in\mathbb{N}} \frac{1}{n!}\left(\sum_{i=0}^n \frac{n!}{i!(n-i)!} a^i b^{n-i}\right) T^n \\ &= \sum_{n\in\mathbb{N}} \frac{(a+b)^n}{n!} T^n \\ &= \exp\left((a+b)T\right).\end{aligned}$$

□

Definition 6.11.14 Let $P = \sum_{n\in\mathbb{N}} a_n T^n \in K[T]$. If $P \neq 0$, we denote by $\deg P$ the greatest index $n \in \mathbb{N}$ such that $a_0 \neq 0$. $A_{\deg(P)}$ is called the **leading coefficient** of P. If $P = 0$, by convention, $\deg P = -\infty$.

Proposition 6.11.15 If P and Q are elements in $K[T]$,

$$\deg(P+Q) \leq \max\{\deg(P), \deg(Q)\},$$

$$\deg(PQ) \leq \deg(P) + \deg(Q).$$

Theorem 6.11.16 Let $P \in K[T], P \neq 0$ Let d be the degree of P and c be the leading coefficient of P. Assume that $c \in K^\times$. For any $F \in K[T]$, there exists a unique $(Q, R) \in K[T]^2$, such that $\deg(R) < d$ and $F = PQ + R$.

Proof By induction on $\deg(F)$. If $\deg(F) < d$, take $Q = 0, R = F$. Suppose that

$$F = a_n T^n + a_{n-1} T^{n-1} + \cdots + a_0,\ n \geq d, a_n \neq 0.$$

Let $G = F - c^{-1} a_n P T^{n-d}$. Then $\deg(G) < n$. Apply the induction hypothesis to G. We get $(Q_1, R) \in K[T]^2, \deg(R) < d$, such that $G = Q_1 P + R$. So $F = P(Q_1 + c^{-1} a_n T^{n-d}) + R$.
"Uniqueness": If $F = PQ_1 + R_1 = PQ_2 + R_2$. Then $P(Q_1 - Q_2) = R_2 - R_1$.

$$\deg(P(Q_1 - Q_2)) = \deg(P) + \deg(Q_1 - Q_2), \deg(R_2 - R_1) < d.$$

So $\deg(Q_1 - Q_2) = \deg(R_2 - R_1) = -\infty$. □

Definition 6.11.17 If the leading coefficient of a non-zero polynomial $F \in K[T]$ is 1, we say that F is **monic**.

Theorem 6.11.18 If K is a field, then $K[T]$ is a principal ideal domain.

Proof We have seen that $K[\![T]\!]$ is a integral domain, so is $K[T]$. Let I be an ideal $K[T]$. If $I = \{0\}$, it is generated by 0. Suppose that $I \neq \{0\}$. There exists a monic $P \in I$ such that $\deg(P)$ is the least among the non-zero polynomials of I. For any $F \in I$, when we write F into the form $F = PQ + R$ with $\deg(R) < \deg(P)$. One has $R \in I$, so $R = 0$. We then get

$$I = K[T]P := \{PQ \mid Q \in K[T]\}.$$

□

Chapter 7

Limit

7.1 Filters

Definition 7.1.1 Let X be a set. We call **filter** on X any non-empty subset $\mathcal{F}$ of $\mathscr{P}(X)$ this satisfies:
(1) $\forall (V_1, V_2) \in \mathcal{F}^2, V_1 \cap V_2 \in \mathcal{F}$.
(2) $\forall V \in \mathcal{F}, \forall W \in \mathscr{P}(X)$, if $V \subseteq W$, then $W \in \mathcal{F}$.

Remark 7.1.2
If $\varnothing \in \mathcal{F}$, then $\mathcal{F} = \mathscr{P}(X)$, we say that $\mathcal{F}$ is degenerate.

Example 7.1.3 If $Y \subseteq X$, then

$$\mathcal{F}_Y := \{V \in \mathscr{P}(X) \mid Y \subseteq V\}$$

is a filter, called the principal filter of Y.

If $\mathcal{F}$ is a non-degenerate filter such that, for any non-degenerate filter $\mathcal{G}$, one has $\mathcal{F} \nsubseteq \mathcal{G}$. We say that $\mathcal{F}$ is an **ultrafilter**.

Proposition 7.1.4 Let I be a non-empty set and $(\mathcal{F}_i)_{i\in I}$ is a family of filters on X, then $\mathcal{F} := \bigcap_{i\in I} \mathcal{F}_i$ is also a filter on X.

Proof
(1) $\forall (V_1, V_2) \in \mathcal{F}^2$, one has

$$\forall i \in I, (V_1, V_2) \in \mathcal{F}_i^2,$$

so $V_1 \cap V_2 \in \mathcal{F}_i$. This leads to $V_1 \cap V_2 \in \mathcal{F}$.
(2) $\forall V \in \mathcal{F}$, one has $\forall i \in I, V \in \mathcal{F}_i$. If $W \in \mathscr{P}(X), W \supseteq V$, then $\forall i \in I, W \in \mathcal{F}_i$. □

Definition 7.1.5 Let S be a subset of $\mathscr{P}(X)$. We denote by $\mathcal{F}_S$ the intersection of all filters containing S. It is thus the least filter containing S. We call it the filter generated by S.

Remark 7.1.6 If $Y \subseteq X$, then the principal filter $\mathcal{F}_Y$ is generated by $\{Y\}$.

Proposition 7.1.7 Let X be a set and S be a non-empty subset of $\mathscr{P}(X)$, then

$$\mathcal{F}_S := \{U \in \mathscr{P}(X) \mid \exists n \in \mathbb{N}_{\geq 1}, \exists (A_1, \dots, A_n) \in S^n, A_1 \cap \cdots \cap A_n \subseteq U\}.$$

Proof Denote by $\mathcal{F}'_S$ the set on the right hand side of the equality. One has $\mathcal{F}'_S \subseteq \mathcal{F}_S$[a] . It remains to check that $\mathcal{F}'_S$ is a filter containing S. By definition, $S \subseteq \mathcal{F}'_S$. If $(U, V) \in \mathcal{F}'^2_S, \exists A_1, \dots, A_n, B_1, \dots, B_n \in S, A_1 \cap \cdots \cap A_n \subseteq U, B_1 \cap \cdots \cap B_n \subseteq V$, so $A_1 \cap \cdots \cap A_n \cap B_1 \cap \cdots \cap B_n \subseteq U \cap V$. If $W \supseteq U$, then $A_1 \cap \cdots \cap A_n \subseteq W$, so $W \in \mathcal{F}'_S$. □

[a] $A_i \in S \subseteq \mathcal{F}_S$. Since $\mathcal{F}_S$ is a filter $U \in \mathcal{F}_S$.

Definition 7.1.8 We say that a subset S of $\mathscr{P}(X)$ is a **filter basis** if, for any $(A, B) \in S \times S$, there exists $C \in S$, such that $C \subseteq A \cap B$.[a]

[a] If $n \in \mathbb{N}_{\geq 1}$ and $(A_1, \dots, A_n) \in S^n, \exists C \in S$ such that $C \subseteq A_1 \cap \cdots \cap A_n$.

Remark 7.1.9 If S is a filter basis, then

$$\mathcal{F}_S = \{U \in \mathscr{P}(X) \mid \exists A \in S, A \subseteq U\}.$$

If S is a subset of $\mathscr{P}(X)$, then

$$\mathcal{B}_S := \{A_1 \cap \cdots \cap A_n \mid n \in \mathbb{N},\ (A_1, \dots, A_n) \in S^n\}$$

is a filter basis containing S. Moreover, $\mathcal{F}_S = \mathcal{F}_{\mathcal{B}_S}$.

Proposition 7.1.10 Let X be a set. Then

$$\mathcal{F} = \{U \in \mathscr{P}(X) \mid X \backslash U \text{ is finite}\}$$

is a filter on X. We call it the **Fréchet filter** of X.

Proof
If $(U,V) \in \mathcal{F}^2, X\backslash(U \cap V) = (X\backslash U) \cup (X\backslash V)$, is finite.
If $U \in \mathcal{F}, W \in \mathscr{P}(X), U \subseteq W$, then $(X\backslash W) \subseteq (X\backslash U)$ is finite. □

Example 7.1.11 Let $I \subseteq \mathbb{N}$ be an infinite set. Let $J \subseteq \mathbb{N}$ be infinite, then $\{I_{\geq j} \mid j \in J\}$ is a filter basis that generates the Fréchet filter of I. $\{I_{\geq j} \mid j \in J\}$ is a totally ordered subset of $\mathscr{P}(I)$, so it is a filter basis. For any $j \in J, I\backslash I_{\geq j} = I_{<j}$ is finite. Let $U \in$ Fréchet filter of I, $I\backslash U$ is finite. There exists $j \in J$ such that $\forall i \in I\backslash U, i < j$. So $I\backslash U \subseteq I_{<j}, U \supseteq I\backslash I_{j<} = I_{\geq j}$.

Example 7.1.12 Let X be a set. We call **pseudometric** on X any mapping

$$\mathrm{d} : X \times X \to \mathbb{R}_{\geq 0}.$$

such that,
(1) $\forall x \in X, \mathrm{d}(x,x) = 0.$
(2) $\forall (x,y) \in X^2, \mathrm{d}(x,y) = \mathrm{d}(y,x).$
(3) (Triangle inequality) $\forall (x,y,z) \in X^3, \mathrm{d}(x,z) \leq \mathrm{d}(x,y) + \mathrm{d}(y,z).$
(X, d) is called the **pseudometric space**. If

$$\forall (x,y) \in X^2, x \neq y \Rightarrow \mathrm{d}(x,y) > 0,$$

then (X, d) is called a **metric space**.
Let (X, d) be a pseudometric space. For any $x \in X$, and $\varepsilon \in \mathbb{R}_{\geq 0}$, we denote by $B(x,\varepsilon)$ the set

$$\{y \in X \mid \mathrm{d}(x,y) < \varepsilon\},$$

called the **open ball** *center at x of radius ε*.
Then

$$\mathcal{V}_x := \{U \in \mathscr{P}(X) \mid \exists \varepsilon \in \mathbb{R}_{>0}, B(x,\varepsilon) \subseteq U\}$$

is a filter, called the **filter of neighborhood** of x.

Proposition 7.1.13 Let $J \subseteq \mathbb{R}_{>0}$ be a non-empty subset such that $\inf J = 0$. Then $\mathcal{B}_J = \{B(x,\varepsilon) \mid \varepsilon \in J\}$ is a filter basis such that $\mathcal{F}_{\mathcal{B}_J} = \mathcal{V}_x$.

Proof $\forall U \in \mathcal{V}_x, \exists \varepsilon \in J, \varepsilon < \delta,$

$$B(x, \varepsilon) \subseteq B(x, \delta) \subseteq U.$$

□

7.2 Order Limit

We fix a partially ordered set $(G, \leq)$ assumed to be order complete.

Example 7.2.1
(1) $\mathbb{R} \cup \{-\infty, +\infty\},\ \forall x \in \mathbb{R}, -\infty < x < +\infty.$
(2) $[0, +\infty]$.
(3) $(\mathscr{P}(\Omega), \subseteq)$.

Definition 7.2.2 Let X be a set and $f : X \longrightarrow G$ be a mapping. For any $U \in \mathscr{P}(X)$, we define

$$f^s(U) := \sup_{x \in U} f(x) = \sup f(U).$$

$$f^i(U) := \inf_{x \in U} f(x) = \inf f(U).$$

If $U \neq \varnothing$, $f^s(U) \geq f^i(U)$. Let $\mathcal{F}$ be a filter on X. We define

$$\limsup_{\mathcal{F}} f := \inf_{U \in \mathcal{F}} f^s(U).$$

$$\liminf_{\mathcal{F}} f := \sup_{U \in \mathcal{F}} f^i(U).$$

They are called the **superior limit** and the **inferior limit** of f along $\mathcal{F}$. If

$$\liminf_{\mathcal{F}} f = \limsup_{\mathcal{F}} f,$$

we say that f has a limit along $\mathcal{F}$, and we denote $\lim_{\mathcal{F}} f$ this value.

Notation 7.2.3 Let $I \subseteq \mathbb{N}$ be an infinite subset. We call sequence in G parametrized by I any element of $G^I = \{(a_n)_{n \in I} \mid \forall n \in I, a_n \in G\}$. If $\mathcal{F}$ is the Fréchet filter on I, then for any $f = (a_n)_{n \in I} \in G^I$, $\limsup_{\mathcal{F}} f$ is denote as

$\limsup_{n\in I, n\to+\infty} a_n$ or as $\limsup_{n\to+\infty} a_n$. Resp. $\liminf$.

Proposition 7.2.4 Let $f : X \longrightarrow G$ be a mapping and $\mathcal{F}$ be a non-degenerate filter. Then

$$\forall (U, V) \in \mathcal{F} \times \mathcal{F}, f^s(U) \geq f^i(V).$$

In particular

$$\limsup_{\mathcal{F}} f \geq \liminf_{\mathcal{F}} f.$$

Proof

$$f^s(U) \geq f^s(U \cap V) \geq f^i(U \cap V) \geq f^i(V).$$

Taking $\inf_{U\in\mathcal{F}}$, we get $\forall V \in \mathcal{F}$, $\limsup_{\mathcal{F}} f \geq f^i(V)$. Taking $\sup_{V\in\mathcal{F}}$, we get $\limsup_{\mathcal{F}} f \geq \liminf_{\mathcal{F}} f$. $\square$

Proposition 7.2.5 Let $f : X \longrightarrow G$ be a mapping, $\mathcal{B}$ be a filter basis on X and $\mathcal{F}$ be the filter generated by $\mathcal{B}$. Then

$$\limsup_{\mathcal{F}} f = \inf_{B\in\mathcal{B}} f^s(B), \ \liminf_{\mathcal{F}} f = \sup_{B\in\mathcal{B}} f^i(B).$$

Proof Since $\mathcal{B} \subseteq \mathcal{F}$, one has

$$\limsup_{\mathcal{F}} f = \inf_{U\in\mathcal{F}} f^s(U) \leq \inf_{B\in\mathcal{B}} f^s(B).$$

For any $U \in \mathcal{F}, \exists A \in \mathcal{B}$ such that $U \supseteq A$. One has

$$f^s(U) \geq f^s(A) \geq \inf_{B\in\mathcal{B}} f^s(B).$$

Taking $\inf_{U\in\mathcal{F}}$, we get

$$\limsup_{\mathcal{F}} f \geq \inf_{B\in\mathcal{B}} f^s(B).$$

$\square$

Consequence: If $I \subseteq \mathbb{N}$ is an infinite subset, $J \subseteq \mathbb{N}$ is another infinite subset, $\forall (a_n)_{n\in I} \in G^I$,

$$\limsup_{n\in I, n\to+\infty} a_n = \inf_{j\in J} \sup_{n\in I_{\geq j}} a_n,$$

$$\liminf_{n\in I, n\to+\infty} a_n = \sup_{j\in J} \inf_{n\in I_{\geq j}} a_n.$$

Example 7.2.6 $a_n = (-1)^n$, $(a_n)_{n\in\mathbb{N}} \in [-\infty, +\infty]^N$,

$$\limsup_{n\to+\infty}(-1)^n = \inf_{j\in 2\mathbb{N}} \sup_{n\geq j}(-1)^n = \inf_{j\in 2\mathbb{N}} 1 = 1.$$

$$\liminf_{n\to+\infty}(-1)^n = -1.$$

Example 7.2.7 $\left(\frac{1}{n}\right)_{n\in\mathbb{N}_{\geq 1}}$,

$$\limsup_{n\to+\infty} \frac{1}{n} = \inf_{j\in\mathbb{N}_{\geq 1}} \sup_{n\geq j} \frac{1}{n} = \inf_{j\in\mathbb{N}_{\geq 1}} \frac{1}{j} = 0,$$

$$\liminf_{n\to+\infty} \frac{1}{n} = \sup_{j\in\mathbb{N}_{\geq 1}} \inf_{n\geq j} \frac{1}{n} = \sup_{j\in\mathbb{N}_{\geq 1}} 0 = 0.$$

Proposition 7.2.8 Let $f, g : X \longrightarrow G$ be mappings and $\mathcal{F}$ be a filter on X. Suppose that there exists $A \in \mathcal{F}$ such that

$$\forall x \in A, f(x) \leq g(x).$$

Then,

$$\limsup_{\mathcal{F}} f \leq \limsup_{\mathcal{F}} g, \ \liminf_{\mathcal{F}} f \leq \liminf_{\mathcal{F}} g.$$

Proof Let

$$\mathcal{B} = \{U \in \mathcal{F} \mid U \subseteq A\}.$$

$\mathcal{B}$ is a filter basis, and $\mathcal{B} \in \mathcal{F}$. For any $V \in \mathcal{F}$, one has $V \cap A \in \mathcal{B}$ and $V \supseteq V \cap A$. So $\mathcal{F}$ is generated by $\mathcal{B}$. For any $B \in \mathcal{B}$, one has $B \subseteq A$ and hence

$$f^s(B) \leq g^s(B), \ f^i(B) \leq g^i(B).$$

So

$$\inf_{B\in\mathcal{B}} f^s(B) \leq \inf_{B\in\mathcal{B}} g^s(B), \ \sup_{B\in\mathcal{B}} f^i(B) \leq \sup_{B\in\mathcal{B}} g^i(B).$$

□

Theorem 7.2.9 (Squeeze Theorem) Let X be a set and $\mathcal{F}$ be a non-degenerate filter on X. Let f, g, h be elements of G^X. Assume that there exists $A \in \mathcal{F}$ such that

$$\forall x \in A, f(x) \leq g(x) \leq h(x).$$

If f and h have limits along $\mathcal{F}$, and

$$\lim_{\mathcal{F}} f = \lim_{\mathcal{F}} h,$$

then, g also has a limit along $\mathcal{F}$, and

$$\lim_{\mathcal{F}} f = \lim_{\mathcal{F}} g = \lim_{\mathcal{F}} h.$$

Proof

$$\lim_{\mathcal{F}} f = \limsup_{\mathcal{F}} f \leq \limsup_{\mathcal{F}} g \leq \limsup_{\mathcal{F}} h = \lim_{\mathcal{F}} h.$$

So

$$\limsup_{\mathcal{F}} g = \lim_{\mathcal{F}} f = \lim_{\mathcal{F}} h.$$

$$\lim_{\mathcal{F}} f = \liminf_{\mathcal{F}} f \leq \liminf_{\mathcal{F}} g \leq \liminf_{\mathcal{F}} h = \lim_{\mathcal{F}} h.$$

So

$$\liminf_{\mathcal{F}} g = \lim_{\mathcal{F}} f = \lim_{\mathcal{F}} h.$$

□

Example 7.2.10 Let $a > 1$. Consider the sequence $\left(\frac{a^n}{n!}\right)_{n\in\mathbb{N}}$. If $n \geq N \geq 2a$, $a \leq \frac{N}{2}$, then

$$0 \leq \frac{a^n}{n!} \leq \frac{a^N}{N!} \cdot \frac{a^{n-N}}{(N+1)\ldots n} \leq \frac{a^N}{N!} \frac{1}{2^{n-N}}.$$

For any $n \geq N$, $0 \leq \frac{a^n}{n!} \leq \frac{(2a)^N}{N!} \cdot \frac{1}{2^n}$. So by squeeze theorem, $\lim\limits_{n\to+\infty} \dfrac{a^n}{n!} = 0$.

Theorem 7.2.11 (Monotone Convergence Theorem) Let I be an infinite subset of $\mathbb{N}$ and $(a_n)_{n\in I} \in G^I$.
(1) If $(a_n)_{n\in I}$ is increasing, then $(a_n)_{n\in I}$ admits $\sup\limits_{n\in I} a_n$ as its limit.

(2) If $(a_n)_{n\in I}$ is decreasing, then $(a_n)_{n\in I}$ admits $\inf\limits_{n\in I} a_n$ as its limit.

Proof
(1) Let $l = \sup_{n\in I} a_n, \forall n \in \mathbb{N},\ a_n \leq l$. So

$$\limsup_{n\to+\infty} a_n \leq \limsup_{n\to+\infty} l = l.$$

$$\forall j \in I,\ \inf_{n\in I_{\geq j}} a_n = a_j,$$

so

$$\liminf_{n\to+\infty} a_n = \sup_{j\in I} \inf_{n\in I_{\geq j}} a_n = \sup_{j\in I} a_j = l.$$

Hence,

$$l = \liminf_{n\to+\infty} a_n \leq \limsup_{n\to+\infty} a_n \leq l.$$

Which means

$$\lim_{n\to+\infty} a_n = l.$$

$\square$

Proposition 7.2.12 Let X be a set and $Y \subseteq X$.
(1) If $\mathcal{F}$ is a filter on X, then

$$\mathcal{F}|_Y := \{U \cap Y \mid U \in \mathcal{F}\}$$

is a filter on Y.
(2) If $\mathcal{B}$ is a filter basis on X, and $\mathcal{F}$ is the filter generated by $\mathcal{B}$, then

$$\mathcal{B}|_Y := \{B \cap Y \mid B \in \mathcal{B}\}$$

is a filter basis generates $\mathcal{F}|_Y$.

Proof
(1) Let U and V be elements of $\mathcal{F}$, one has

$$(U \cap Y) \cap (V \cap Y) = (U \cap V) \cap Y \in \mathcal{F}|_Y .$$

Let $U \in \mathcal{F}, W \subseteq Y, U \cap Y \subseteq W$. Let $V = U \cup W \in \mathcal{F}$.

$$Y \cap V = (U \cap Y) \cup (W \cap Y) = W.$$

Hence $W \in \mathcal{F}|_Y$.
(2) Let B_1, B_2 be elements of $\mathcal{B}$, then $\exists A \in B, A \subseteq B_1 \cap B_2$. Thus

$$A \cap Y \subseteq (B_1 \cap Y) \cap (B_2 \cap Y).$$

So $\mathcal{B}|_Y$ is a filter basis. Moreover, $\mathcal{B}|_Y \subseteq \mathcal{F}|_Y$. Let $U \in \mathcal{F}, \exists B \in \mathcal{B}$ such that $B \subseteq U$. Thus

$$B \cap Y \subseteq U \cap Y.$$

So $U \cap Y$ contains an element of $\mathcal{B}|_Y$. □

Example 7.2.13 Let $I \subseteq \mathbb{N}$ be an infinite subset, and $(a_n)_{n\in I} \in G^I$. If $J \subseteq I$ is an infinite subset, $\mathcal{F}$ be the filter on I, then $\mathcal{F}|_J$ is the Fréchet filter on J. $(a_n)_{n\in J}$ is called a subsequence of $(a_n)_{n\in I}$.

Proposition 7.2.14 Let $f : X \longrightarrow G$ be a mapping, $\mathcal{F}$ be a filter on X, $Y \subseteq X$. Then
(1)

$$\limsup_{\mathcal{F}|_Y} f|_Y \leq \limsup_{\mathcal{F}} f,$$

$$\liminf_{\mathcal{F}|_Y} f|_Y \geq \liminf_{\mathcal{F}} f.$$

(2) Suppose that $\mathcal{F}|_Y$ is non-degenerate and f has a limit along $\mathcal{F}$, then $f|_Y$ has a limit along $\mathcal{F}|_Y$ and

$$\lim_{\mathcal{F}} f = \lim_{\mathcal{F}|_Y} f|_Y.$$

(3) If $Y \in \mathcal{F}$, then

$$\limsup_{\mathcal{F}|_Y} = \limsup_{\mathcal{F}} f,$$

$$\liminf_{\mathcal{F}|_Y} = \liminf_{\mathcal{F}} f.$$

Proof
$\forall U \in \mathcal{F}, f^s(U \cap Y) \leq f^s(U)$. So

$$\limsup_{\mathcal{F}|_Y} f|_Y = \inf_{U\in\mathcal{F}} f^s(U \cap Y) \leq \inf_{U\in\mathcal{F}} f^s(U) = \limsup_{\mathcal{F}} f.$$

(2)

$$\lim_{\mathcal{F}} f = \limsup_{\mathcal{F}} f \geq \limsup_{\mathcal{F}|_Y} f|_Y \geq \liminf_{\mathcal{F}|_Y} f|_Y \geq \liminf_{\mathcal{F}} f = \lim_{\mathcal{F}} f.$$

(3) $\mathcal{F}|_Y$ is a filter basis that generates $\mathcal{F}$ if $Y \in \mathcal{F}$,

$$\limsup_{\mathcal{F}|_Y} f|_Y = \inf_{V \in \mathcal{F}|_Y} f^s(U) = \inf_{U \in \mathcal{F}} f^s(U) = \limsup_{\mathcal{F}} f.$$

□

Theorem 7.2.15 (Bolzano-Weierstrass) Suppose that G is totally ordered. Let $I \subseteq \mathbb{N}$ be an infinite subset and $(a_n)_{n \in I}$ be a sequence in G.
(1) There exists an infinite subset J_1 such that $(a_n)_{n \in J_1}$ is monotone and admits $\limsup_{n \in I, n \to +\infty} a_n$ as its limit.
(2) There exists an infinite subset J_2 such that $(a_n)_{n \in J_2}$ is monotone and admits $\liminf_{n \in I, n \to +\infty} a_n$ as its limit.

Proof
(1) Let
$$J = \{n \in I \mid \forall m \in I_{\geq n}, a_m \leq a_n\}.$$
If J is infinite, $(a_n)_{n \in J}$ is decreasing. Hence it admits
$$\alpha := \inf_{n \in J} a_n$$
as its limit. For any $n \in J$, $\sup_{m \in I_{\geq n}} a_m = a_n$. So
$$\limsup_{n \in I, n \to +\infty} a_n = \inf_{n \in J} \sup_{m \in I_{\geq n}} a_m = \alpha.$$

Suppose that J is finite. Pick $n_0 \in I$ such that $\forall j \in J, j < n_0$. We construct in a recursive way a strictly increasing sequence $(n_k)_{k \in \mathbb{N}}$ in I as follows: Suppose $n_0 < n_1 < \cdots < n_k$ have been chosen. Since G is totally ordered, there exists $i \in I$ such that $n_0 \leq i \leq n_k$ and
$$a_i = \max\{a_j \mid j \in I, n_0 \leq j \leq n_k\}.$$
Since $i \notin J$, there exists $n_{k+1} \in I, n_{k+1} > i$ such that
$$a_{n_{k+1}} > a_i.$$
Note that $n_{k+1} > n_k$. Let
$$J_1 = \{n_k \mid k \in \mathbb{N}\},$$

$(a_n)_{n \in J_1}$ is increasing, hence it admits

$$\beta := \sup_{n \in J} a_n$$

as its limit. For any $j \in I$ such that $j \geq n_0$, there exists $k \in \mathbb{N}$ such that $j \leq n_k$. Thus $a_j \leq a_{n_{k+1}} \leq \beta$. So $\limsup_{n \in I, n \to +\infty} a_n \leq \beta$. Moreover, since $J_1 \subseteq I$,

$$\beta = \lim_{n \in J_1, n \to +\infty} a_n = \limsup_{n \in J_1, n \to +\infty} a_n \leq \limsup_{n \in I, n \to +\infty} a_n.$$

Therefore,

$$\beta = \limsup_{n \in I, n \to +\infty} a_n.$$

□

7.3 Partially Ordered Groups

Definition 7.3.1 Let $(G, *)$ be a group, and $\leq$ be a partial order on G. If

$$\forall (a, b, c) \in G^3, a < b \Rightarrow a * c < b * c \text{ and } c * a < c * b,$$

we say that $(G, *, \leq)$ is a **partially ordered group**. If in addition $\leq$ is a total order, we say that $(G, *, \leq)$ is a **totally ordered group**. (Resp. semigroup, monoid.)

Example 7.3.2 (1) $(\mathbb{R}, +, \leq)$. (2) $(\mathbb{R}_{>0}, \cdot, \leq)$. (3) $(\mathbb{N} \setminus \{0\}, \cdot, |)$.

Remark 7.3.3
(1) If $(G, *)$ is a partially ordered group, then

$$\forall (a, b, c) \in G^3, a \leq b \Rightarrow a * c \leq b * c, \ c * a \leq c * b.$$

(2) $(G, \hat{*}, \leq)$ is a partially ordered group.
Resp. semigroup, monoid.

Proposition 7.3.4
Let $(G, *, \leq)$ be a partially ordered semigroup. Let $(a_1, a_2, b_1, b_2) \in G^4$.
(1) If $a_1 \leq a_2, \ b_1 \leq b_2$, then $a_1 * b_1 \leq a_2 * b_2$.
(2) If $a_1 < a_2, \ b_1 \leq b_2$, then $a_1 * b_1 < a_2 * b_2$.

(3) If $a_1 \le a_2,\ b_1 < b_2$, then $a_1 * b_1 < a_2 * b_2$.

Proof
(1) $a_1 * b_1 \le a_2 * b_1 \le a_2 * b_2$
(2),(3) At least one of the above inequality is strict. □

Proposition 7.3.5 Let $(G, *, \le)$ be a partially ordered semigroup, $(x, y, a) \in G^3$. Assume that, either $\le$ is a total order, or $(G, *)$ is a monoid and $a \in G^\times$. Then the following conditions are equivalent:
(1) $x \le y$.
(2) $x * a \le y * a$.
(3) $a * x \le a * y$.

Proof By definition, $(1) \Rightarrow (2)$, $(1) \Rightarrow (3)$. Assume that $x * a \le y * a$. If $(G, *)$ is a monoid and $a \in G^\times$, then

$$x = (x * a) * \iota(a) \le (y * a) * \iota(a) = y.$$

Suppose that $\le$ is a total order. If $x \not\le y$, then $x > y$ and $x * a > y * a$, contradiction. □

Corollary 7.3.6 A totally ordered semigroup satisfies the left and right cancellation laws.

Proof Let $(G, *, \le)$ be a totally ordered semigroup. Let $(x, y, a) \in G^3$ such that $x * a = y * a$. Then

$$x * a \le y * a, y * a \le x * a.$$

Hence $x \le y$ and $y \le x$. □

Proposition 7.3.7 Let $(G, *, \le)$ be a partially ordered monoid. Then,

$$\iota : G^\times \longrightarrow G^\times$$

is strictly decreasing.

Proof Let $(x, y) \in G^\times \times G^\times$ such that $x < y$. Then

$$e = \iota(x) * x < \iota(x) * y,$$

where e is the neutral element of $(G, *)$. Thus

$$e * \iota(y) < \iota(x) * y * \iota(y).$$

That is $\iota(y) < \iota(x)$. □

Proposition 7.3.8 Let $(G, *, \leq)$ be a totally ordered group, and e be the neutral element of $(G, *)$. If $G \neq \{e\}$, then G has neither a greatest element nor a least element.

Proof Suppose that $(G, \leq)$ has a greatest element β. We first show by contradiction that $\beta \neq e$. Suppose that $e = \max G$. Pick $x \in G, x \neq e$. Then $x < e$. Thus $\iota(x) > \iota(e)$. Contradiction. If $\beta > e, \beta * \beta > e * \beta = \beta$, contradiction, too. □

7.4 Enhancement

Definition 7.4.1 Let $(S, *, \leq)$ be a partially ordered semigroup. Suppose that $(S, \leq)$ has no greatest element and has no least element. Let $\perp$ and $\top$ be formal elements and let

$$\bar{S} = S \cup \{\perp, \top\}.$$

We extend $\leq$ to $\bar{S}$ by letting $\perp < x < \top, \forall x \in S$. We extend $*$ to a mapping

$$\left(\bar{S} \times \bar{S}\right) \backslash \{(\perp, \top), (\top, \perp)\} \longrightarrow \bar{S},$$

such that

$$\forall x \in S \cup \{\top\}, x * \top = \top * x = \top.$$

$$\forall x \in S \cup \{\perp\}, x * \perp = \perp * x = \perp.$$

$\top * \perp$ and $\perp * \top$ are NOT DEFINED. $(\bar{S}, *, \leq)$ is called the **enhancement** of $(S, *, \leq)$. If A and B are subset of $\bar{S}$, we denote by $A * B$ the set

$$\{x * y \mid (x, y) \in A \times B \,, \{x, y\} \neq \{\perp, \top\}\}.$$

Example 7.4.2
(1) $(\mathbb{R}, +, \leq), \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$.
(2) $(\mathbb{R}_{>0}, \cdot, \leq), \bar{\mathbb{R}}_{>0} = \mathbb{R}_{>0} \cup \{0, +\infty\}$.

Remark 7.4.3
(1) $\forall (a, b) \in \bar{S} \times \bar{S}$, $a * b$ is defined if and only if $b * a$ is defined.
(2) If $*$ is commutative, and $a * b$ is defined, then $a * b = b * a$.

Definition 7.4.4 Let $a_0, \dots, a_n$ be elements of $\bar{S}$, if $a_0 * \dots * a_n$ is defined, and $(a_0, \dots, a_{n-1}) * a_n$ is also defined, then we let $a_0 * \dots * a_n = (a_0 * \dots * a_{n-1}) * a_n$.

Proposition 7.4.5 Let $a_0, \dots, a_n$ be elements of $\bar{S}$. For any $i \in \{0, \dots, n\}$, $a_0 * \dots * a_{i-1} * (a_i * a_{i+1}) * \dots * a_n$ is defined if and only if $a_0 * \dots * a_n$ is defined. Moreover, $a_0 * \dots * a_n = a_0 * \dots (a_{i-1} * a_i) * \dots * a_n$.

Proof Both terms are defined is and only if

$$\{\top, \bot\} \nsubseteq \{a_0, \dots, a_n\}.$$

If $\bot \in \{a_0, \dots, a_n\}$, then both terms are equal to $\bot$. If $\top \in \{a_0, \dots, a_n\}$, then both terms are equal to $\top$. □

Proposition 7.4.6 Let $(S, *, \leq)$ be a partially ordered semigroup. Let $(a, b) \in \bar{S} \times \bar{S}$. If $a < b$, then for any $c \in S$, $a * c < b * c, c * a < c * b$.

Proof If $\{a, b\} \subseteq S$, this follows from the definition of a partially ordered semigroup. If $a = \bot$, then $b > \bot$. $a * c = c * a = \bot$. $b * c \neq \bot, c * b \neq \bot$. So $a * c < b * c, c * a < c * b$. If $\{a, b\} \subseteq S$ and $a \neq \bot$, then $b = \top$, $b * c = c * b = \top$. $a * c \neq \top, c * a \neq \top$. So $a * c < b * c$, $c * a < c * b$. □

Proposition 7.4.7
Let $(S, *, \leq)$ be a partially ordered semigroup and $(x, y, a, b) \in \bar{S}^4$.
(1) If $x < a$ and $y < b$, then $x * y$ and $a * b$ are defined, and $x * y < a * b$.
(2) If $x \leq a, y \leq b$ and $x * y$ and $a * b$ are defined, $x * y \leq a * b$.

Proof
(1) Since $x < a, y < b, \top \notin \{x,y\}, \bot \notin \{a,b\}$. So $x * y$ and $a * b$ are defined. If $\top \in \{a,b\}$, then $a * b = \top$. Since $\top \notin \{x,y\}, x * y \neq \top$, so $x * y < a * b$. If $\bot \in \{x,y\}$, then $x * y = \bot$. Since $\bot \notin \{a,b\}, a * b \neq \bot$. So $x * y < a * b$. If $\top \notin \{a,b\}, \bot \notin \{x,y\}$,, then $\{x,y,a,b\} \subseteq S$. So $x * y < x * b < a * b$.
(2) If $\top \in \{x,y\}$, then $\top \in \{a,b\}$, so $x * y = \top = a * b$. If $\bot \in \{x,y\}$, then $\bot \in \{a,b\}$, so $x * y = \bot = a * b$. If $\top \in \{x,y\}$, then $x * y = \bot \leq a * b$. If $\bot \in \{x,y\}$, then $x * y \geq \top = a * b$. If $\{x,y,a,b\} \subseteq S$, then $x * y \leq a * y \leq a * b$.
□

Proposition 7.4.8 Let $(S, *, \leq)$ be a partially ordered monoid and $e \in S$ be the neutral element. Let $(a,b) \in \bar{S} \times \bar{S}$, with $a \in S^{\times} \cup \{\bot, \top\}$. Then the following conditions are equivalent.
(1) $a < b$.
(2) $\iota(a) * b$ is defined and $e < \iota(a) * b$. (Where $\iota(\bot) = \top, \iota(\top) = \bot$.)
(3) $b * \iota(a)$ is defined and $e < b * \iota(a)$.

Proof Suppose that $a < b$. Then a and b cannot be both $\top$ or be both $\bot$. Hence $\{\iota(a), b\} \neq \{\bot, \top\}$. Therefore, $\iota(a) * b$ and

$$b * \iota(a)$$

are defined. If $\{a,b\} \subseteq S$, then $e = \iota(a) * a < \iota(a) * b$, $e = a * \iota(a) < b * \iota(a)$. If $a = \bot$, then $\iota(a) = \top$ and $\iota(a) * b = b * \iota(a) = \top$. So $e < \iota(a) * b, e < b * \iota(a)$. If $b = \top$, then $\iota(a) * b = b * \iota(a) = \top > e$.
Assume (2). $\iota(a) * b$ is defined and $e < \iota(a) * b$. If $a \in S$,

$$a = a * e < a * (\iota(a) * b) = (a * \iota(a)) * b = e * b = b.$$

If $a = \bot, \iota(a) = \top, b \neq \bot$, so $a < b$. If $a = \top, \iota(a) = \bot, \iota(a) * b = \bot < e$, contradiction. □

Corollary 7.4.9 Let $(S, *, \leq)$ be a partially ordered monoid and $(a,b) \in (S^{\times} \cup \{\bot, \top\})^2$. Then $a < b$ if and only if $\iota(a) > \iota(b)$.

Proof If $a < b$,then $\iota(a) * b$ is defined and

$$e < \iota(a) * b = \iota(a) * \iota(\iota(b)).$$

So $\iota(b) < \iota(a)$. □

Lemma 7.4.10 Let $(S, *, \leq)$ be a partially ordered monoid. Let $A \subseteq \bar{S}, b \in S^{\times} \cup \{\perp, \top\}$. Then the following statements hold:
(1) If $\sup(A) * b$ is defined, then $A * \{b\}$ has a supremum in $\bar{S}$, and

$$\sup(A * \{b\}) = \sup(A) * b.$$

(2) If $\inf(A) * b$ is defined, then $A * \{b\}$ has a infimum in $\bar{S}$, and

$$\inf(A * \{b\}) = \inf(A) * b.$$

(3) If $b * \sup(A)$ is defined, then $A * \{b\}$ has a supremum in $\bar{S}$, and

$$\sup(\{b\} * A) = b * \sup(A).$$

(4) If $b * \inf(A)$ is defined, then $A * \{b\}$ has a infimum in $\bar{S}$, and

$$\inf(\{b\} * A) = b * \inf(A).$$

Proof
(1) Suppose that $b = \perp$, then $\sup(A) \neq \top$, $\sup(A) * b = \perp$. $A * \{b\} \subseteq \{\perp\}$, so $\sup(A * \{b\}) = \perp$. Suppose that $b = \top$, then $\sup(A) \neq \perp$, $A \neq \varnothing$, $\sup(A) * b = \top$. So $\sup(A * \{b\}) = \top = \sup(A) * \top$. We suppose that $b \in S^{\times}$. $\forall a \in A,\ A \leq \sup(A)$, so $a * b \leq \sup(A) * b$. This means that $\sup(A) * b$ is an upper bound of $A * \{b\}$. Let M be an upper bound of $A * \{b\}$. For any $a \in A$, $a \times b \in A * \{b\}$, so $a * b \leq M$. Hence,

$$a = (a * b) * \iota(b) \leq M * \iota(b).$$

We then deduce $\sup(A) \leq M * \iota(b)$. Hence $\sup(A) * b \leq M * \iota(b) * b$. Therefore, $\sup(A) * b$ is the supremum of $A * \{b\}$. □

Remark 7.4.11 Consider

$$S = \{0\} \cup [2, 3[\cup [4, +\infty[\subseteq \mathbb{R}.$$

$A = [2, 3[, \sup(A) = 4, A + \{2\} = [4, 5[, \sup(A + \{2\}) = 5, \sup(A) + 2 = 6.$

Theorem 7.4.12 Let $(S, *, \leq)$ be a partially ordered group. Let A and B be subsets of $\bar{S}$.

(1) If $\sup(A) * \sup(B)$ is defined, then $A * B$ has a supremum in $\bar{S}$ and

$$\sup(A * B) = \sup(A) * \sup(B).$$

(2) If $\inf(A) * \inf(B)$ is defined, then $A * B$ has a infimum in $\bar{S}$ and

$$\inf(A * B) = \inf(A) * \inf(B).$$

Proof For any $(a, b) \in A \times B$, if $a * b$ id defined, then

$$a * b \leq \sup(A) * \sup(B).$$

So $\sup(A) * \sup(B)$ is an upper bound of $A * B$. If $\perp \in \{\sup(A), \sup(B)\}$, then $A * B$ has $\perp$ as an upper bound. So $\sup(A * B) = \perp = \sup(A) * \sup(B)$. We suppose that $\perp \notin \{\sup(A) * \sup(B)\}$. Thus $A \backslash \{\perp\} \neq \varnothing$, $B \backslash \{\perp\} \neq \varnothing$. Suppose that $\sup(A) = \top$. Take $b \in B \backslash \{\perp\}$.

$$\sup(A * B) \geq \sup(A) * \{b\} = \sup(A) * b = \top.$$

So $\sup(A * B) = \top = \sup(A) * \sup(B)$. Similarly, if $\sup(B) = \top$, then

$$\sup(A * B) = \top = \sup(A) * \sup(B).$$

Suppose that $\top \notin \{\sup(A), \sup(B)\}$. For any $b \in B$, $\sup(A) * b$ is defined since $\sup(A) \in S$. Hence

$$\sup(A) * \{b\} = \sup(A) * b.$$

$$A * B = \bigcup_{b \in B} A * \{b\},$$

$$\{\sup(A) * b \mid b \in B\} = \{\sup(A)\} * B.$$

By the lemma, $\{\sup(A)\} * B$ has a supremum, which is $\sup(A) * \sup(B)$. So $\sup(A * B)$ exists, and is equal to

$$\{\sup(A * \{b\}) \mid b \in B\} = \sup(A) * \sup(B).$$

□

Corollary 7.4.13 Let $(S, *, \leq)$ be a partially ordered group. Let $f, g : X \longrightarrow \bar{S}$ be two mappings. Let

$$Y = \{x \in X \mid f(x) * g(x) \text{ is defined }\}.$$

Let

$$f * g : Y \longrightarrow \bar{S},$$

$$y \longmapsto f(y) * g(y).$$

(1) If $(\sup f) * (\sup g)$ is defined, and $f * g$ has a supremum, then

$$\sup(f * g) \leq \sup(f) * \sup(g).$$

(2) If $(\inf f) * (\inf g)$ is defined, and $f * g$ has a infimum, then

$$\inf(f * g) \geq \inf(f) * \inf(g).$$

Proof Let $A = f(X), B = g(X)$. By the theorem, $A * B$ has a supremum, and

$$\sup(A * B) = \sup(A) * \sup(B).$$

Let

$$C = (f * g)(Y) = \{f(y) * g(y) \mid y \in Y\}.$$

One has

$$C \subseteq A * B = \{f(x) * g(y) \mid (x, y) \in X \times X,\ f(x) * g(y) \text{ is defined }\}.$$

So $\sup(C) \leq \sup(A * B)$. $\square$

Theorem 7.4.14 Let $(S, *, \leq)$ be a partially ordered group. We suppose that $\bar{S}$ is order complete. Let X be a set and $f, g : X \longrightarrow \bar{S}$ be mappings. Let $\mathcal{F}$ be a filter on X that is non-degenerate. Suppose that $\forall x \in X, f(x) * g(x)$ is defined. Then

$$\limsup_{\mathcal{F}}(f * g) \leq \limsup_{\mathcal{F}} f * \limsup_{\mathcal{F}} g,$$

$$\limsup_{\mathcal{F}}(f * g) \geq \limsup_{\mathcal{F}} f * \liminf_{\mathcal{F}} g,$$

$$\liminf_{\mathcal{F}}(f * g) \geq \liminf_{\mathcal{F}} f * \liminf_{\mathcal{F}} g,$$

$$\liminf_{\mathcal{F}}(f * g) \leq \liminf_{\mathcal{F}} f * \limsup_{\mathcal{F}} g.$$

Provided that the term on the right hand side is defined.

Proof

(1) $\forall U \in \mathcal{F}$,

$$(f * g)^s(U) \leq f^s(U) * g^s(U).$$

Provided that $f^s(U) * g^s(U)$ is defined. If $\limsup_{\mathcal{F}} f * \limsup_{\mathcal{F}} g$ is defined, then

$$\limsup_{\mathcal{F}}(f * g) = \inf_{U \in \mathcal{F}}(f * g)^s(U) \leq \inf_{U \in \mathcal{F}}\left[f^s(U) * g^s(U)\right] = l.$$

$$\begin{aligned}\limsup_{\mathcal{F}} f * \limsup_{\mathcal{F}} g &= \left(\inf_{U \in \mathcal{F}} f^s(U)\right) * \left(\inf_{V \in \mathcal{F}} g^s(V)\right) \\ &= \inf_{(U,V) \in \mathcal{F} \times \mathcal{F},\text{ defined}} \left(f^s(U) * g^s(V)\right)\end{aligned}$$

If $(U, V) \in \mathcal{F} \times \mathcal{F}$ is such that $f^s(U) * g^s(V)$ is defined, then

$$f^s(U) * g^s(V) \geq f^s(U \cap V) * g^s(U \cap V) \geq l$$

provided that $f^s(U \cap V) * g^s(U \cap V)$ is defined. If $f^s(U \cap V) * g^s(U \cap V)$ is not defined, then $\top \in \{f^s(U), g^s(V)\}$, so that

$$f^s(U) * g^s(V) = \top \geq l.$$

Therefore,

$$\limsup_{\mathcal{F}} f * \limsup_{\mathcal{F}} g \geq l \geq \limsup_{\mathcal{F}} f * g.$$

(2)

$$\limsup_{\mathcal{F}} f * g \geq \limsup_{\mathcal{F}} f * \liminf_{\mathcal{F}} g.$$

Let $U \in \mathcal{F}$. Suppose that $(f * g)^s(U) \neq \top$. $\forall V \in \mathcal{F}$, one has $\forall x \in U \cap V$,

$$(f * g)^s(U \cap V) \geq f(x) * g(x) \geq f(x) * g^i(U \cap V) \geq f(x) * g^i(V).$$

So

$$(f * g)^s(U) \geq f^s(U \cap V) * g^i(V),$$

provided that $f^s(U \cap V) * g^i(V)$ is defined. Taking the infimum which respect to U, we obtain

$$\limsup_{\mathcal{F}} f * g \geq \limsup_{\mathcal{F}} f * g^i(V),$$

provided that $\limsup_{\mathcal{F}} f * g^i(V)$ is defined. Taking the supremum with respect to V, we obtain

$$\limsup_{\mathcal{F}} f * g \geq \limsup_{\mathcal{F}} f * \liminf_{\mathcal{F}} g.$$

□

Corollary 7.4.15 Let $(S, *, \leq)$ be a partially ordered group, such that $\bar{S}$ is order complete. Let $f, g : X \longrightarrow \bar{S}$ be mappings such that $\forall x \in X, f(x) * g(x)$ is defined. Let $\mathcal{F}$ be a non-degenerate filter on X. Assume that g has a limit along $\mathcal{F}$.
(1)

$$\limsup_{\mathcal{F}} f * g = \limsup_{\mathcal{F}} f * \lim_{\mathcal{F}} g.$$

(2)

$$\liminf_{\mathcal{F}} f * g = \liminf_{\mathcal{F}} f * \lim_{\mathcal{F}} g.$$

Provided that the term on the right hand side is defined.

Proof

$$\limsup_{\mathcal{F}} f * g \leq \limsup_{\mathcal{F}} f * \limsup_{\mathcal{F}} g = \limsup_{\mathcal{F}} f * \lim_{\mathcal{F}} g.$$

$$\limsup_{\mathcal{F}} f * g \geq \limsup_{\mathcal{F}} f * \liminf_{\mathcal{F}} g = \limsup_{\mathcal{F}} f * \lim_{\mathcal{F}} g.$$

□

Example 7.4.16
(1)

$$\limsup_{n \to +\infty} \left((-1)^n + \frac{1}{n} \right) = \limsup_{n \to +\infty} (-1)^n = 1.$$

(2) For $a > 1$, let $a = 1 + b$. $a^n = (1 + b)^n \geq nb$. so

$$0 \leq \frac{\sqrt{n}}{a^n} \leq \frac{1}{b\sqrt{n}},$$

$$\lim_{n \to +\infty} \frac{\sqrt{n}}{a^n} = 0.$$

$$\forall k \in \mathbb{N}_{\geq 1}, \lim_{n \to +\infty} \left(\frac{\sqrt{n}}{a^n} \right)^k = 0.$$

Proposition 7.4.17 Let $(G, \leq)$ be a totally ordered set that is order complete. Let X be a set and $f, g : X \longrightarrow G$ be mappings. We denote by

$$f \wedge g : X \longrightarrow G, x \longmapsto \min\{f(x), g(x)\},$$

$$f \vee g : X \longrightarrow G, x \longmapsto \max\{f(x), g(x)\}.$$

Let $\mathcal{F}$ be a filter on X that is non-degenerate. Then

$$\limsup_{\mathcal{F}} (f \vee g) = \max\{\limsup_{\mathcal{F}} f, \limsup_{\mathcal{F}} g\},$$

$$\liminf_{\mathcal{F}} (f \wedge g) = \min\{\liminf_{\mathcal{F}} f, \liminf_{\mathcal{F}} g\}.$$

If f and g have limit along $\mathcal{F}$, so do $f \vee g$ and $f \wedge g$ and

$$\limsup_{\mathcal{F}} (f \vee g), \limsup_{\mathcal{F}} (f \wedge g).$$

Proof Let

$$\mathcal{B}_1 = \{A \in \mathcal{F} \mid f^s(A) \geq g^s(A)\},$$

$$\mathcal{B}_2 = \mathcal{F} \backslash \mathcal{B}_1 = \{A \in \mathcal{F} \mid f^s(A) < g^s(A)\}.$$

Case 1. $\forall U \in \mathcal{F}, \exists A \in \mathcal{B}_1, A \subseteq U$.

$$\limsup_{\mathcal{F}}(f \vee g) = \inf_{\mathcal{F}}(f \vee g)^s(U) = \inf_{A \in \mathcal{B}_1}(f \vee g)^s(A) = \inf_{A \in \mathcal{B}_1} f^s(A) = \limsup_{\mathcal{F}} f.$$

Case 2. $\exists W \in \mathcal{F}, \forall A \in \mathcal{B}_1, A \nsubseteq W$. If $\mathcal{B}_1 = \varnothing, \mathcal{B}_2 = \mathcal{F}$. $(U \in \mathcal{F}, U \in \mathcal{B}_2)$ If $A \in \mathcal{B}_1, \forall U \in \mathcal{F}, U \subseteq W, U \cap A \notin \mathcal{B}_1$, so $U \cap A \in \mathcal{B}_2$ and $U \cap A \subseteq U$. (U contains an elements in $\mathcal{B}_2$.)

$$\limsup_{\mathcal{F}}(f \vee g) = \inf_{U \in \mathcal{F}|_W}(f \vee g)^s(U) = \inf_{A \in \mathcal{B}_2}(f \vee g)^s(A) = \inf_{A \in \mathcal{B}_2} g^s(A) = \limsup_{\mathcal{F}} g.$$

Hence,

$$\limsup_{\mathcal{F}}(f \vee g) \leq \max\{\limsup_{\mathcal{F}} f, \limsup_{\mathcal{F}} g\}.$$

Moreover, $f \vee g \geq f, f \vee g \geq g$, so

$$\limsup_{\mathcal{F}}(f \vee g) \geq \max\{\limsup_{\mathcal{F}} f, \limsup_{\mathcal{F}} g\}.$$

If $\lim_{\mathcal{F}} f$ and $\lim_{\mathcal{F}} g$ exists, then

$$\begin{aligned}\liminf_{\mathcal{F}}(f \wedge g) &\geq \min\{\liminf_{\mathcal{F}} f, \liminf_{\mathcal{F}} g\} \\ &= \max\{\lim_{\mathcal{F}} f, \lim_{\mathcal{F}} g\} = \limsup_{\mathcal{F}}(f \vee g).\end{aligned}$$

$\square$

Proposition 7.4.18 Let $(S, *, \leq)$ be a partially orders group. For any $A \subseteq S$, let

$$\iota(A) := \{\iota(a) \mid a \in A\}.$$

If A has an supremum in $\bar{S}$ then $\iota(A)$ has an infimum in $\bar{S}$ and

$$\inf(\iota(A)) = \iota(\sup(A)).$$

Resp. infimum, supremum.

Proof
$\forall a \in A, a \leq \sup(A)$, so $\iota(a) \geq \iota(\sup(A))$. Hence $\iota(\sup(A))$ is a lower bound of $\iota(A)$. Let m be a lower bound of $\iota(A)$. For any $a \in A$, $\iota(a) \geq m, \iota(m) \geq a$. Hence $\iota(m) \geq \sup(A), \iota(\sup(A)) \geq m$. □

Corollary 7.4.19 Let $(S, *, \leq)$ be a partially orders group. $f : X \longrightarrow \bar{S}$ be a mapping, and $\mathcal{F}$ be a filter on X. Let

$$\iota(f) : X \longrightarrow \bar{S},\ x \longmapsto \iota(f(x)).$$

Assume that $\bar{S}$ is order complete. Then

$$\limsup_{\mathcal{F}} \iota(f) = \iota(\liminf_{\mathcal{F}} f),$$

$$\liminf_{\mathcal{F}} \iota(f) = \iota(\limsup_{\mathcal{F}} f).$$

Proof

$$\limsup_{\mathcal{F}} \iota(f) = \inf_{U \in \mathcal{F}} = \inf_{U \in \mathcal{F}} \iota\left(\inf_{x \in u} f(x)\right) = \iota\left(\sup_{U \in \mathcal{F}} \inf_{x \in U} f(x)\right) = \iota\left(\liminf_{\mathcal{F}} f\right).$$

□

7.5 Absolute Values

We fix a totally ordered abelian group $(R, +, \leq)$.

Remark 7.5.1 $\forall a \in R$, either $a = 0$, or $a > 0$, or $a < 0$. We add two final elements $-\infty, +\infty$ to R to construct the enhancement of $(R, +, \leq)$.

Definition 7.5.2 For any $x \in \bar{R}$, we let

$$|x| := \max\{x, -x\}.$$

By definition, $|-x| = |x|$.

Proposition 7.5.3 For any $(a, b) \in \bar{R} \times \bar{R}$, if $a + b$ is defined, then

$$|a + b| \leq |a| + |b| .$$

Proof If $\{-\infty, +\infty\} \cap \{a, b\} \neq \varnothing$, then

$$|a| + |b| = +\infty \leq |a + b|$$

if $a + b$ is defined. Suppose that $\{a, b\} \subseteq R$,

$$a + b \leq |a| + |b| ,$$

$$-(a + b) = (-a) + (-b) \leq |a| + |b| ,$$

so,

$$|a + b| \leq |a| + |b| .$$

$\square$

Corollary 7.5.4 For any $(a, b) \in \bar{R} \times \bar{R}$,

$$||a| - |b|| \leq |a - b| , ||a| - |b|| \leq |a + b| ,$$

provided that the terms on both sides are defined.

Proof If $\{-\infty, +\infty\} \cap \{a, b\} \neq \varnothing$, $|a - b| = +\infty$. Suppose that $\{a, b\} \subseteq R$.

$$|a| - |b| = |(a - b) + b| - |b| \leq |a - b| + |b| - |b| = |a - b| .$$

Similarly,

$$|b| - |a| \leq |b - a| = |a - b| .$$

So,

$$||a| - |b|| \leq |a - b| .$$

$\square$

Theorem 7.5.5 Let X be a set, $\mathcal{F}$ be a non-degenerate filter on X, $f : X \longrightarrow R$ be a mapping and $l \in R$. The following statements are equivalent.
(1) f admits l along $\mathcal{F}$.
(2) $\limsup_{\mathcal{F}} |f - l| = 0$, where $f - l : X \longrightarrow R, x \longmapsto f(x) - l$. Moreover, there conditions imply.
(3) $\forall \varepsilon \in R_{>0}, \exists U \in \mathcal{F}, \forall x \in U, |f(x) - l| < \varepsilon$. The converse is true when $\inf R_{>0} = 0$.

Proof Note that

$$\limsup_{\mathcal{F}} (f - l) = \left(\limsup_{\mathcal{F}} f\right) - l,$$
$$\liminf_{\mathcal{F}} (f - l) = \left(\liminf_{\mathcal{F}} f\right) - l.$$

(1)⇒(2): (1) gives $\limsup_{\mathcal{F}} f = \liminf_{\mathcal{F}} f = l$. So

$$\limsup_{\mathcal{F}} (f - l) = \liminf_{\mathcal{F}} (f - l) = 0.$$

So,

$$\begin{aligned}\limsup_{\mathcal{F}} |f - l| &= \limsup_{\mathcal{F}} (f - l) \vee (l - f)\\ &= \max\{\limsup_{\mathcal{F}}(f - l), \limsup_{\mathcal{F}}(l - f)\}\\ &= 0.\end{aligned}$$

(2)⇒(1): $\max\{\limsup_{\mathcal{F}}(f - l), \limsup_{\mathcal{F}}(l - f)\} = 0$. So

$$\left(\limsup_{\mathcal{F}} f\right) - l \leq 0, \quad l - \liminf_{\mathcal{F}} f \leq 0.$$

Hence,

$$l \leq \liminf_{\mathcal{F}} f \leq \limsup_{\mathcal{F}} f \leq l.$$

(2)⇒(3): One has $\inf\limits_{U \in \mathcal{F}} \sup\limits_{x \in U} |f(x) - l| = 0$. For any $\varepsilon \in R_{>0}$, ε is not a lower bound of

$$\left\{\sup_{x \in U} |f(x) - l| \mid U \in \mathcal{F}\right\}.$$

So, $\exists U \in \mathcal{F}, \sup\limits_{x \in U} |f(x) - l| < \varepsilon$. So $\forall x \in U, |f(x) - l| < \varepsilon$.
(3)⇒(2): Under the hypothesis, $\inf R_{>0} = 0$.

$$\forall \varepsilon \in R_{>0}, \exists U \in \mathcal{F}, \forall x \in U, |f(x) - l| < \varepsilon.$$
$$0 \leq \limsup_{\mathcal{F}} |f - l| \leq \varepsilon.$$

Taking the infimum with respect to $\varepsilon \in R_{>0}$, we obtain $\limsup_{\mathcal{F}} |f - l| = 0$. □

Definition 7.5.6 Let $A \subseteq R$. If $\exists M \in R_{>0}$ such that $\forall a \in A, |a| \leq M$,we say that A is **bounded**. If $f : X \longrightarrow R$ is such that $f(X)$ is bounded, we say that f is bounded.

Corollary 7.5.7 Let X be a set, $\mathcal{F}$ be a non-degenerate filter on X, and $f : X \longrightarrow R$ be a mapping. If f has a limit along $\mathcal{F}$ and $\lim_{\mathcal{F}} f \in R$, then there exists $U \in \mathcal{F}$ such that $f|_U$ is bounded.

Proof If $R = \{0\}$, then f is constant, so bounded. Suppose that there exists $\varepsilon \in R_{>0}$,

$$\exists U \in \mathcal{F}, \forall x \in U, |f(x) - l| < \varepsilon,$$

where $l = \lim f$. So $|f(x)| \leq |f(x) - l| + |l| \leq \varepsilon + |l|$, on U. $\square$

7.6 Ordered Rings

Definition 7.6.1 Let $(R, +, \cdot)$ be a non-zero unitary ring and $(R, +, \leq)$ be a partially ordered group and

$$\forall (a, b) \in R_{>0} \times R_{>0}, ab > 0,$$

then we say that $(R, +, \cdot, \leq)$ is an **partially ordered ring**.
Let $\bar{R} = R \cup \{-\infty, +\infty\}$ be an enhancement of R. We extend partially the multiplication on $\bar{R}$ as follows

$$\forall x \in \bar{R}_{>0},\ x(+\infty) = (+\infty)x = +\infty,\ x(-\infty) = (-\infty)x = -\infty.$$

$$\forall x \in \bar{R}_{<0},\ x(+\infty) = (+\infty)x = (-\infty),\ x(-\infty) = (-\infty)x = +\infty.$$

$$\cancel{0 \cdot \infty}$$

are NOT defined.

$$(+\infty)^{-1} = (-\infty)^{-1} = 0.$$

We fix a partially ordered ring $(R, +, \cdot, \leq)$.

(1) If $a \geq 0, b \geq 0$, then $ab \geq 0$.
(2) $a < b \Rightarrow -b < -a$.
(3) $a \leq b \Rightarrow -b \leq -a$.
(4) If $\lambda > 0,\ a > b$ then $\lambda a > \lambda b,\ a\lambda > b\lambda,\ (-\lambda)a < (-\lambda)b,\ a(-\lambda) < b(-\lambda)$.
(5) If $\lambda > 0,\ a \geq b$, then $\lambda a \geq \lambda b,\ a\lambda \geq b\lambda,\ (-\lambda)a \leq (-\lambda)b,\ a(-\lambda) \leq b(-\lambda)$.

Proposition 7.6.2 Let $(R, +, \cdot, \leq)$ be a totally ordered unitary ring.
(1) $\forall a \in R,\ a \neq 0 \Rightarrow a^2 > 0$.
(2) $\forall a \in R^{\times},\ a > 0 \Rightarrow a^{-1} > 0, a < 0 \Rightarrow a^{-1} < 0$.

Proof
(1) Since $\leq$ is a total order, either $a > 0,\ a^2 = a \cdot a > 0$, or $a < 0,\ a^2 = (-a) \cdot (-a) > 0$.
(2) $a^2 > 0$. If $a^{-1} < 0$, then $a = a^{-1} \cdot a^2 < 0$. If $a^{-1} > 0,\ a = a^{-1} \cdot a^2 > 0$. □

Proposition 7.6.3 Let $(R, +, \cdot, \leq)$ be a totally ordered unitary ring. The unique homomorphism of rings

$$f : \mathbb{Z} \longrightarrow R,$$

$$n \longmapsto \begin{cases} n1_R,\ n > 0 \\ -(-n)1_R,\ n < 0 \end{cases}$$

is strictly increasing. In particular, it is injective.

Proof We first prove that $1_R > 0_R$. Since $(R, \leq)$ is totally ordered, either $1_R > 0_R$, or $1_R < 0_R$. If $1_R < 0_R$, then $-1_R > 0_R$. So $(-1_R)(-1_R) = 1_R > 0$, contradiction. Let $(a, b) \in \mathbb{R}^2$, such that $n < m$. Then $\exists k \in \mathbb{N}_{>0}, m = n + k$. So

$$f(m) = f(n) + f(k) = f(n) + 1_R + \cdots + 1_R > f(n).$$

□

Proposition 7.6.4 Let $(R, +, \cdot, \leq)$ be a non-zero totally ordered unitary ring. If $f : \mathbb{Q} \longrightarrow R$ is a ring homomorphism, then f is strictly increasing.

Proof Let $\frac{a}{b}$ and $\frac{c}{d}$ be element in $\mathbb{Q}$, where $(b, d) \in \mathbb{Z}^2_{>0},\ (a, c) \in \mathbb{Z}^2$, such that

$\frac{a}{b} > \frac{c}{d}$. Then $f\left(\frac{a}{b}\right) = \frac{f(a)}{f(b)}$, $f\left(\frac{c}{d}\right) = \frac{f(c)}{f(d)}$. So,

$$f\left(\frac{a}{b}\right) - f\left(\frac{c}{d}\right) = \frac{f(a)f(d) - f(b)f(c)}{f(b)f(d)} = \frac{f(ad-bc)}{f(b)f(d)} > 0.$$

$\square$

Proposition 7.6.5 Let $(R, +, \cdot, \leq)$ be a totally ordered unitary ring.

$$\forall (a,b) \in R^2,\ |ab| = |a|\,|b|$$

Proof $|a| \in \{a, -a\},\ |b| \in \{b, -b\}, |a||b| \in \{ab, -ab\}$. Moreover, $|a||b| \geq 0$. So $|a||b| = |ab|$. $\square$

Theorem 7.6.6 Let $(R, +, \cdot, \leq)$ be a totally ordered unitary ring. Let X be a set and $\mathcal{F}$ be a non-degenerate filter on X. Let $f, g : X \longrightarrow R$ be mappings. Assume that $\bar{R}$ is order complete, f and g have limits (in $\bar{R}$) along $\mathcal{F}$, and $a := \lim\limits_{\mathcal{F}} f,\ b := \lim\limits_{\mathcal{F}} g$. If ab is defined, then the mapping $fg : X \longrightarrow R, x \longmapsto f(x)g(x)$ has a limit along $\mathcal{F}$, and $\lim\limits_{\mathcal{F}} fg = ab$.

Proof Suppose ab is defined. Since f and g have limits along $\mathcal{F}$ in $\bar{R}$, there exists $U \in \mathcal{F},\ \exists M \in \bar{R}_{\geq 0}$ such that $\forall x \in U,\ |f(x)| \leq M$. Therefore,

$$\begin{aligned}\forall x \in U,\ &|f(x)g(x) - ab|\\ &= |f(x)g(x) - f(x)b + f(x)b - ab|\\ &\leq |f(x)\left(g(x) - b\right)| + |\left(f(x) - a\right)b|\\ &= |f(x)|\,|g(x) - b| + |f(x) - a|\,|b|\\ &\leq M\,|g(x) - b| + |b|\,|f(x) - a|\,.\end{aligned}$$

Hence,

$$\limsup_{\mathcal{F}} |fg - ab| \leq M \limsup |g(x) - b| + |b| \limsup_{\mathcal{F}} |f(x) - a| = 0.$$

Therefore,

$$\lim_{\mathcal{F}} fg = ab.$$

$\square$

Chapter 8

Topology

8.1 Topological spaces

Proposition 8.1.1 Let X be a set and for any $x \in X$, let $\mathcal{G}_x$ be a filter contained in the principal filter of $\{x\}$ ($\forall U \in \mathcal{G}_x,\ x \in U$). Denote by $\mathscr{T}$ the set

$$\{U \in \mathscr{P}(X) \mid \forall x \in U, U \in \mathcal{G}_x\}.$$

Then the following conditions are satisfied.
(1) $\{\varnothing, X\} \subseteq \mathscr{T}$.
(2) If $(U_1, U_2) \in \mathscr{T}^2$, then $U_1 \cap U_2 \in \mathscr{T}$.
(3) If I is a set and $(U_i)_{i \in I} \in \mathscr{T}^I$, then $\bigcup_{i \in I} U_i \in \mathscr{T}$.
Moreover, $\forall x \in X,\ \mathcal{B}_x = \{U \in \mathscr{T} \mid x \in U\}$ is a filter basis contained in $\mathcal{G}_x$. It generates $\mathcal{G}_x$ if the following condition is satisfied:

$$\forall U \in \mathcal{G}_x,\ \exists V \in \mathcal{G}_x,\ V \subseteq U \text{ and } \forall y \in V,\ V \in \mathcal{G}_y.$$

Proof
(1) $\varnothing \in \mathscr{T}, X \in \bigcap_{x \in X} \mathcal{G}_x$.
(2) $\forall x \in U_1 \cap U_2,\ U_1 \in \mathcal{G}_x,\ U_2 \in \mathcal{G}_x$, so $U_1 \cap U_2 \in \mathcal{G}_x$.
(3) Let $U = \bigcup_{i \in I} U_i$. $\forall x \in U,\ \exists i \in I,\ x \in U_i$, so $U_i \in \mathcal{G}_x$. Since $U \supseteq U_i$, so $U \in \mathcal{G}_x$.

$$\mathcal{B}_x := \{U \in \mathscr{T} \mid x \in U\}.$$

If $U \in \mathcal{B}_x$, then $x \in U$, so $U \in \mathcal{G}_x$. Hence $\mathcal{B}_x \subseteq \mathcal{G}_x$. If $(U, V) \in \mathcal{B}_x^2$, then $U \cap V \in \mathscr{T}$, and $x \in U \cap V$. So $U \cap V \in \mathcal{B}_x$. So $\mathcal{B}_x$ is a filter basis. Suppose the condition is satisfied. For any $U \in \mathcal{G}_x$, $\exists V \in \mathcal{G}_x \cap \mathscr{T}$, such that $V \subseteq U$. Note that $V \in \mathcal{B}_x$, so $\mathcal{G}_x$ is generated by $\mathcal{B}_x$. $\square$

Definition 8.1.2 Let X be a set. We call **topology** on X any subset $\mathscr{T}$ of $\mathscr{P}(X)$ that satisfies the following conditions:
(1) $\{\varnothing, X\} \subseteq \mathscr{T}$.
(2) If $(U_1, U_2) \in \mathscr{T}^2$, then $U_1 \cap U_2 \in \mathscr{T}$.
(3) For any set I, $\forall\,(U_i)_{i\in I} \in \mathscr{T}^I$, then $\bigcup_{i\in I} U_i \in \mathscr{T}$.
$(X, \mathscr{T})$ is called a **topological space**.
If $\forall x \in X,\ \mathcal{B}_x$ is a filter basis of X contained in the principal filter of $\{x\}$, then

$$\mathscr{T} = \{U \in \mathscr{P}(X) \mid \forall x \in U,\ \exists V_x \in \mathcal{B}_x,\ V_x \subseteq U\}$$

is a topology on X, called the **topology generated by** $(\mathcal{B}_x)_{x\in X}$. More generally, if $\forall x \in X,\ S_x$ is a subset of the principal filter of $\{x\}$ and $\mathcal{G}_x$ is the filter generated by S_x, then we say that

$$\mathscr{T} = \{U \in \mathscr{P}(X) \mid \forall x \in U,\ U \in \mathcal{G}_x\}$$

is the topology generated by $(S_x)_{x\in X}$.

Example 8.1.3
(1) Let $\mathcal{G}_x = \{X\}$. The topology generated by $(\mathcal{G}_x)_{x\in X}$ is $\{\varnothing, X\}$, called the **trivial topology** on X.
(2) Let $\mathcal{G}_x = \mathcal{F}_{\{x\}}$ be the principal filter. The topology generated by $(\mathcal{G}_x)_{x\in X}$ is $\mathscr{P}(X)$. This topology is called the **discrete topology** on X.
(3) Let (X, d) be a semimetric space.

$$\mathrm{d} : X \times X \longrightarrow \mathbb{R}_{\geq 0},\ \mathrm{d}(x,y) = \mathrm{d}(y,x),\ \mathrm{d}(x,z) \leq \mathrm{d}(x,y) + \mathrm{d}(y,z),\ \mathrm{d}(x,x) = 0.$$

$\forall \varepsilon > 0,\ \forall x \in X$, let $B(x,\varepsilon) = \{y \in X \mid \mathrm{d}(x,y) < \varepsilon\}$, $\{B(x,\varepsilon) \mid \varepsilon \in \mathbb{R}_{>0}\} =: \mathcal{B}_x$ is a filter basis on X, $\mathcal{B}_x$ is contained in the principal filter of $\{x\}$. The topology

$$\mathscr{T} = \{U \in \mathscr{P}(X) \mid \forall x \in U,\ \exists \varepsilon \in \mathbb{R}_{>0},\ B(x,\varepsilon) \subseteq U\}$$

is called the **topology induced by the semimetric** d.
(4) Let $(G, \leq)$ be a totally ordered set $\forall x \in G$, let $S_x = \{G_{>a} \mid a < x\} \cup \{G_{<b} \mid x < b\}$

Proposition 8.1.4 $\forall x \in X,\ \forall \varepsilon \in \mathbb{R}_{>0}, B(x,\varepsilon) \in \mathscr{T}$.

Proof $\forall y \in B(x,\varepsilon),\ \mathrm{d}(x,y) < \varepsilon$. Let $r = \varepsilon - \mathrm{d}(x,y) > 0$, we claim that $B(y,r) \subseteq B(x,\varepsilon)$. Let $z \in B(y,r),\ \mathrm{d}(y,z) < r$. Hence,

$$\mathrm{d}(x,z) \leq \mathrm{d}(x,y) + \mathrm{d}(y,z) < \mathrm{d}(x,y) + r = \mathrm{d}(x,y) + \varepsilon - \mathrm{d}(x,y) = \varepsilon.$$

□

Remark 8.1.5 On $\mathbb{R}$, one has a metric

$$\mathrm{d} : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}_{\geq 0},$$

$$(a,b) \longmapsto |a-b|.$$

$$B(x,\varepsilon) = \,]x-\varepsilon, x+\varepsilon[.$$

$$\mathcal{B}_x = \{]x-\varepsilon, x+\varepsilon[\ \mid \varepsilon \in \mathbb{R}_{>0}\}.$$

Let $\mathscr{T}_{\mathrm{d}}$ be the topology generated by $(\mathcal{B}_x)_{x\in\mathbb{R}}$. Let $\mathscr{T}$ be the order topology generated by $(S_x)_{x\in\mathbb{R}}$, where

$$S_x := \{\mathbb{R}_{>a} \mid a < x\} \cup \{\mathbb{R}_{<b} \mid x < b\}.$$

Proposition 8.1.6 For any $x \in \mathbb{R}$, $\mathcal{F}(\mathcal{B}_x) = \mathcal{F}(S_x)$.

Proof $\forall \varepsilon > 0,\ ]x-\varepsilon, x+\varepsilon[\ = \mathbb{R}_{<x+\varepsilon} \cap \mathbb{R}_{>x-\varepsilon} \in \mathcal{F}(S_x)$. So $\mathcal{F}(\mathcal{B}_x) \subseteq \mathcal{F}(S_x)$.

$\forall a \in \mathbb{R},\ a < x,\ \mathbb{R}_{>a} \supseteq\ ]a, 2x-a[\ = \,]x-(x-a), x+(x-a)[,\ \mathbb{R}_{>a} \in \mathcal{F}(\mathcal{B}_x)$.

$$\forall b \in \mathbb{R}, b > x, \mathbb{R}_{<b} \supseteq\ ]2x-b, b[\ = \,]x+(b-x), x+(b-x)[,$$

So, $\mathbb{R}_{<b} \subseteq \mathcal{F}(\mathcal{B}_x)$. Hence $S_x \subseteq \mathcal{F}(\mathcal{B}_x)$, which leads to $\mathcal{F}(S_x) \subseteq \mathcal{F}(\mathcal{B}_x)$. □

Definition 8.1.7 Let $(X,\mathscr{T})$ be a topological space. For any $x \in X$ and any $V \in \mathscr{P}(X)$, if there exists $U \in \mathscr{T}$ such that $x \in U \subseteq V$, then we say that V is a **neighborhood of** x. We call **open subset** of X any subset of X that belongs to $\mathscr{T}$. If $U \in \mathscr{T}$, such that $x \in U$, we say that U is an **open neighborhood of** x. We denote by $\mathcal{V}_x(\mathscr{T})$ the set of all neighborhoods of x.

Proposition 8.1.8 $\mathcal{V}_x(\mathscr{T})$ is a filter on X contained in the principal filter of $\{x\}$. Moreover, the topology generated by $(\mathcal{V}_x(\mathscr{T}))_{x\in X}$ identifies with $\mathscr{T}$.

Proof
(1) If $(V_1, V_2) \in \mathcal{V}_x(\mathscr{T})^2$, $\exists (U_1, U_2) \in \mathscr{T}^2$, such that $x \in U_1 \subseteq V_1$, $x \in U_2 \subseteq V_2$. Hence, $x \in U_1 \cap U_2 \subseteq V_1 \cap V_2$, so $V_1 \cap V_2 \in \mathcal{V}_x(\mathscr{T})$.
(2) If $V \in \mathcal{V}_x(\mathscr{T})$, $W \in \mathscr{P}(X)$, $V \subseteq W$. $\exists U \in \mathscr{T}$, $x \in U \subseteq V \subseteq W$, so $W \in \mathcal{V}_x(\mathscr{T})$. Let $\mathscr{T}'$ be the topology generated by $(\mathcal{V}_x(\mathscr{T}))_{x\in X}$. By definition,

$$\mathscr{T}' = \{U \subseteq X \mid \forall x \in U, U \in \mathcal{V}_x(\mathscr{T})\}.$$

For any $U \in \mathscr{T}$, $\forall x \in U$, U is a open neighborhood of x, so $U \in \mathscr{T}'$. Let $U \in \mathscr{T}'$, $\forall x \in U$, $\exists V_2 \in \mathscr{T}$, $x \in V_x \subseteq U$.

$$U = \bigcup_{x\in U} \{x\} \subseteq \bigcup_{x\in U} V_x \subseteq U.$$

$$U = \bigcup_{x\in U} V_x \in \mathscr{T}.$$

□

Proposition 8.1.9 Let X be a set, $(\mathscr{T}_i)_{i\in I}$ be a family of topologies on X. Then

$$\mathscr{T} = \bigcap_{i\in I} \mathscr{T}_i$$

is a topology on X.

Proof
(1) $\forall i \in I$, $\{\varnothing, X\} \subseteq \mathscr{T}_i$, so $\{\varnothing, X\} \subseteq \mathscr{T}$.
(2) If $(U_1, U_2) \in \mathscr{T}^2$, then for any $i \in I$, $U_1 \cap U_2 \in \mathscr{T}_i$, so $U_1 \cap U_2 \in \bigcap_{i\in I} \mathscr{T}_i$.
(3) For any set J and any $(U_j)_{j\in J} \in \mathscr{T}^J$, one has $\forall i \in I$, $\forall j \in J$, $U_j \in \mathscr{T}_i$, so

$$\bigcup_{j\in J} U_j \in \mathscr{T}_i.$$

Therefore,

$$\bigcup_{j\in J} U_j \in \bigcap_{i\in I} \mathscr{T}_i.$$

□

Definition 8.1.10 Let S be a subset of $\mathscr{P}(X)$, we denote by $\mathscr{T}_S$ the intersection of all topologies containing S, we call it the topology generated by S.

Definition 8.1.11
Let $\mathcal{B}$ be a subset of $\mathscr{P}(X)$, we say that $\mathcal{B}$ is a **topological basis** if:
(1) $X = \bigcup_{V \in \mathcal{B}} V$.
(2) $\forall (U, V) \in \mathcal{B} \times \mathcal{B},\ \forall x \in U \cap V,\ \exists W_x \in \mathcal{B},\ x \in W_x \subseteq U \cap V$.

Definition 8.1.12 Let X be a set and $\mathscr{T}_1$ and $\mathscr{T}_2$ be two topologies on X. If $\mathscr{T}_1 \subseteq \mathscr{T}_2$, we say that $\mathscr{T}_1$ is coarser than $\mathscr{T}_2$ and $\mathscr{T}_2$ is finer than $\mathscr{T}_1$.
If $S \subseteq \mathscr{P}(X)$, we denote by $\mathscr{T}_S$ the intersection of all topology containing S. It is the coarsest topology containing S.

Proposition 8.1.13 Let S be s subset of $\mathscr{P}(X)$. Let

$$\mathcal{B}_S := \{X\} \cup \left\{ \bigcap_{i=1}^{n} A_i \mid n \in \mathbb{N}_{\geq 1},\ (A_1, \ldots, A_n) \in S^n \right\},$$

then, $\mathcal{B}_S$ is a topological basis on X. Moreover, $\mathscr{T}_S = \mathscr{T}_{\mathcal{B}_S}$.

Proof Since $X \in \mathcal{B}_S$, $\bigcup_{V \in \mathcal{B}_S} V = X$. Let $(U, V) \in \mathcal{B}_S \times \mathcal{B}_S$. If $U = X$, then $U \cap V = V \in \mathcal{B}_S$. Similarly, if $V = X$, then $U \cap V = U \in \mathcal{B}_S$.
If $U = A_1 \cap \cdots \cap A_n$, $V = B_1 \cap \cdots \cap B_m$, then $\{A_1, \ldots, A_n, B_1, \ldots, B_m\} \subseteq \mathcal{B}_S$.

$$U \cap V = A_1 \cap \cdots \cap A_n \cap B_1 \cap \cdots \cap B_m \in \mathcal{B}_S.$$

Since $S \subseteq \mathcal{B}_S \subseteq \mathscr{T}_{\mathcal{B}_S}$, so $\mathscr{T}_S \subseteq \mathscr{T}_{\mathcal{B}_S}$, $X \in \mathscr{T}_S$. If $(A_1, \ldots, A_n) \in S^n$, then $(A_1, \ldots, A_n) \in \mathscr{T}_{\mathcal{B}_S}$. So $A_1 \cap \cdots \cap A_n \in \mathscr{T}_S$. Hence $\mathcal{B}_S \subseteq \mathscr{T}_S$. Therefore, $\mathscr{T}_{\mathcal{B}_S} \subseteq \mathscr{T}_S$, so $\mathscr{T}_{\mathcal{B}_S} = \mathscr{T}_S$. $\square$

Proposition 8.1.14 Let $\mathcal{B}$ be a topological basis on a set X. Then

$$\mathscr{T}_{\mathcal{B}} = \left\{ U \in \mathscr{P}(X) \mid \exists \text{ a set } I \text{ and } (V_i)_{i \in I} \in \mathcal{B}^I, U = \bigcup_{i \in I} V_i \right\}.$$

Proof We denote by $\mathscr{T}$ the set

$$\{U \in \mathscr{P}(X) \mid U \text{ can be written as as the union of a family sets in } \mathcal{B}\}.$$

By definition, $\mathcal{B} \subseteq \mathscr{T} \subseteq \mathscr{T}_{\mathcal{B}}$. It remains to check that $\mathscr{T}$ is a topology.
By definition, $X \in \mathscr{T}$, $\varnothing \in \mathscr{T}$. Moreover, the union of a family of elements of $\mathscr{T}$ remains in $\mathscr{T}$. Let $U = \bigcup_{i \in I} U_i$ and $V = \bigcup_{j \in J} V_j$ be elements of $\mathscr{T}$, where, $U_i \in \mathcal{B}, V_j \in \mathcal{B}$. Then

$$U \cap V = \bigcup_{(i,j) \in I \times J} (U_i \cap V_j)\,.$$

For any $x \in U_i \cap V_j$, $\exists W_x^{(i,j)} \in \mathcal{B},\ x \in W_x^{(i,j)} \subseteq U_i \cap V_j$. $U_i \cap V_j = \bigcup_{x \in U_i \cap V_j} W_x^{(i,j)}$, so

$$U \cap V = \bigcup_{(i,j) \in I \times J} \bigcup_{x \in U_i \cap V_j} W_x^{(i,j)}.$$

□

8.2 Convergence

We fix a topology space $(E, \mathscr{T})$, $l \in E$ and $S \subseteq \mathscr{P}(E)$ that generates the filter $\mathcal{V}_l(\mathscr{T})$ of all neighborhood of l.

Definition 8.2.1 Let $f : X \longrightarrow Y$ be a mapping. If $\mathcal{F}$ is a filter on X, we denote by $f_*(\mathcal{F})$ the set $\{B \subseteq Y \mid f^{-1}(B) \in \mathcal{F}\}$.

Proposition 8.2.2 $f_*(\mathcal{F})$ is a filter on Y.

Proof Let $(B_1, B_2) \in f_*(\mathcal{F})$,

$$f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2) \in \mathcal{F}.$$

Let $B \in f_*(\mathcal{F}),\ C \supseteq B.\ f^{-1}(C) \supseteq f^{-1}(B) \in \mathcal{F}$, so $f^{-1}(C) \in \mathcal{F}$. □

Proposition 8.2.3 Let $f : X \longrightarrow Y,\ g : Y \longrightarrow Z$, be mappings. $\mathcal{F}$ be a filter on X. Then

$$(g \circ f)_* (\mathcal{F}) = g_*(f_*(\mathcal{F})).$$

Proof

$$\begin{aligned}(g\circ f)_*(\mathcal{F}) &= \{C\subseteq Z \mid (g\circ f)^{-1}(C)\in\mathcal{F}\}\\ &= \{C\subseteq Z \mid f^{-1}(g^{-1}(C))\in\mathcal{F}\}\\ &= \{C\subseteq Z \mid g^{-1}(C)\in f_*(\mathcal{F})\}\\ &= g_*(f_*(\mathcal{F})).\end{aligned}$$

□

Proposition 8.2.4 Let $\mathcal{B}$ be a filter basis in X, $f : X \longrightarrow Y$ be a mapping and $\mathcal{F}$ be the filter generated by $\mathcal{B}$. Then $f(\mathcal{B}) : \{f(U) \mid U \in \mathcal{B}\}$ is a filter basis on Y and $f_*(\mathcal{F})$ is the filter generated by $f(\mathcal{B})$.

Proof Let U and V be elements of $\mathcal{B}$. Then $\exists W \in \mathcal{B}$, $W \subseteq U \cap V$. Hence $f(W) \subseteq f(U\cap V) \subseteq f(U)\cap f(V)$. Moreover, for any $U\in\mathcal{B}$, $U \subseteq f^{-1}(f(U))$. So $f^{-1}(f(U)) \in \mathcal{F}$. Therefore, $f(\mathcal{B}) \subseteq f_*(\mathcal{F})$. Let $A \in f_*(\mathcal{F})$. Then, $f^{-1}(A) \in \mathcal{F}$. So $\exists V \in \mathcal{B}$, $V \subseteq f^{-1}(A)$. Hence $f(V) \subseteq A$. Therefore, $f_*(\mathcal{F})$ is a filter basis generated by $f_*(\mathcal{B})$ □

Definition 8.2.5 Let $f : X \longrightarrow E$ be a mapping, $\mathcal{F}$ be a non-degenerate filter on X. If $f_*(\mathcal{F}) \supseteq \mathcal{V}_l(\mathscr{T})$, we say that f **converges** to l along $\mathcal{F}$.

$$\lim_{\mathcal{F}} f = l$$

denotes "f converges to l along $\mathcal{F}$".
This condition is equivalent to

$$\forall V \in S_l,\ f^{-1}(V) \in \mathcal{F}.$$

If $\mathcal{B}$ is a filter basis which generates $\mathcal{F}$. This condition is also

$$\forall V \in S_l,\ \exists U \in \mathcal{B},\ f(U) \subseteq V.$$

Example 8.2.6
(1) Let $I \subseteq \mathbb{N}$ be an infinite subset and $x = (x_n)_{n\in I} \in \mathbb{N}^I$. Let $\mathcal{F}$ be the Fréchet filter on I. If x converges to l along $\mathcal{F}$, or equivalently,

$$\forall V \in S_l,\ \exists N \in \mathbb{N},\ \forall n \in I_{\geq \mathbb{N}},\ x_n \in V.$$

We say that the sequence $(x_n)_{n\in I}$ **converges to** l when n tends to the infinity, denote as

$$\lim_{n\to+\infty} x_n = l.$$

(2) Let $(X, \mathscr{T}_X)$ be a topological space. Let $Y \subseteq X$ be a subset of X, $p \in X$. Let

$$\mathcal{F} = \mathcal{V}_p(\mathscr{T}_X)|_Y = \{V \cap Y \mid V \in \mathcal{V}_p(\mathscr{T}_X)\}.$$

Assume that $\mathcal{F}$ is non-degenerate. Let $f : X \longrightarrow E$ be a mapping. If f converges to l along $\mathcal{F}$, we say that $f(x)$ converges to l when $x \in Y$ tends to p, denoted as

$$\lim_{x\in Y,\ x\to p} f(x) = l.$$

This condition is equivalent to:

$$\forall V \in S_l,\ \exists U \text{ a open neighborhood of } p,\ \forall x \in U \cap Y,\ f(x) \in V.$$

In general, if $g : Y \longrightarrow E$ is a mapping such that

$$\forall V \in S_l,\ \exists U \text{ a open neighborhood of } p,\ \forall x \in U \cap Y,\ g(x) \in V,$$

then we say $g(x)$ converges to l when $x \in Y$ tends to p, denoted as

$$\lim_{x\in Y,\ x\to p} g(x) = l.$$

Remark 8.2.7 If (E, d) is a semimetric space, $\mathscr{T}$ is the semimetric topology. Then condition in (1) becomes:

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \in I_{\geq N},\ \mathrm{d}(x_n, l) < \varepsilon.$$

The conditions in (2) becomes:

$$\forall \varepsilon > 0,\ \exists U \text{ a open neighborhood of } p,\ \forall x \in U \cap Y,\ \mathrm{d}(g(x), l) < \varepsilon.$$

If furthermore, $(X, \mathscr{T}_X)$ is a semimetric space with semimetric d_X. The condition becomes

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in Y,\ \mathrm{d}_X(x, p) < \delta \Rightarrow \mathrm{d}(g(x), l) < \varepsilon.$$

Example 8.2.8
(3) Consider $X = \mathbb{R}$. Let $Y \subseteq X$. Consider the filter $\mathcal{F}$ generated by $\{\mathbb{R}_{>M},\ M \in \mathbb{R}\}$. Suppose that $\mathcal{F}|_Y$ is non-degenerate. Let $g : Y \longrightarrow E$. If $\lim\limits_{\mathcal{F}|_Y} g = l$, we say that $g(x)$ converges to l when x tends to $+\infty$, denoted as

$$\lim_{x\in Y,\ x\to+\infty} g(x) = l.$$

This condition is

$$\forall V \in S_l,\ \exists M \in \mathbb{R}_{>0},\ \forall x \in Y,\ x > M \Rightarrow g(x) \in V.$$

If (E, d) is a metric space, it becomes:

$$\forall \varepsilon > 0,\ \exists M \in \mathbb{R}_{>0},\ \forall x \in Y,\ x > M \Rightarrow \mathrm{d}(g(x), l) < \varepsilon.$$

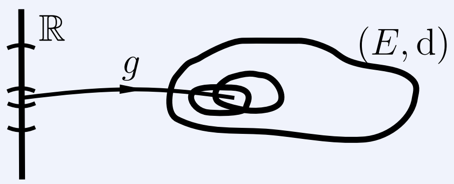


Figure 8.2.1: Example on $\mathbb{R}$

Example 8.2.9 Let $(G, \leq)$ be a totally ordered set, $\mathcal{F}$ be the ordered topology on G. It is generated by $\{G_{>a} \mid a \in G\} \cup \{G_{<b} \mid b \in G\}$. If $l \in G$, then $\mathcal{V}_x(\mathscr{T})$ is generated by

$$S_l := \{G_{>a} \mid a < l\} \cup \{G_{<b} \mid l < b\}.$$

Assume that $(G, \leq)$ is order complete. Let $f : X \longrightarrow G$ be a mapping and $\mathcal{F}$ be a non-degenerate filter on X.
(1) Assume that f converges to l along $\mathcal{F}$. $\forall a < l$, $U_a := f^{-1}(G_{>a}) \in \mathcal{F}$. $\forall x \in U_a, f(x) > a$. So $\liminf_{\mathcal{F}} f \geq a$. If $\sup(G_{<l}) = l$, then $\liminf_{\mathcal{F}} f \geq l$. If $\sup(G_{<l}) < l$, we denote $a = \sup(G_{<l})$. $\forall a \in U_a, f(x) \geq l$. So $\liminf_{\mathcal{F}} f \geq l$. Similarly, $\liminf_{\mathcal{F}} f \leq l$. So f admits l as its limit.
(2) Assume that $\limsup_{\mathcal{F}} f = \liminf_{\mathcal{F}} f = l$.

$$\liminf_{\mathcal{F}} f = l \Rightarrow \sup_{U\in\mathcal{F}} f^i(U) = l, \forall a < l, \exists U \in \mathcal{F}, f^i(U) > a, f^{-1}(G_{>a}) \in \mathcal{F}.$$

$$\limsup_{\mathcal{F}} f = l \Rightarrow \forall b > l, f^{-1}(G_{<b}) \in \mathcal{F}.$$

Therefore, f converges to l along $\mathcal{F}$.

8.3 Continuity

Definition 8.3.1 Let $(X, \mathscr{T}_X)$ and $(Y, \mathscr{T}_Y)$ be topological spaces, f be a function from X to Y, and $p \in \mathrm{Dom}(f)$. If for any neighborhood U of $f(p)$, there exists a neighborhood V of p such that

$$f(V) \subseteq U,$$

then we say that the function f is **continuous** at the point p.
If f is continuous at any $p \in \mathrm{Dom}(f)$, then we say that f is **continuous**.

Remark 8.3.2
(1) The continuity of f at p is equivalent to:

$$\lim_{\substack{x \in \mathrm{Dom}(f) \\ x \to p}} f(x) = f(p),$$

namely, f converges to $f(p)$ when x tends to p.
(2) Let $\mathcal{V}_{f(p)}(\mathscr{T}_Y)$ and $\mathcal{V}_p(\mathscr{T}_X)$ be filters of neighborhoods of $f(p)$ and p respectively. Let $\mathcal{B}_p$ be a filter basis that generates $\mathcal{V}_p(\mathscr{T}_X)$. Let $S_{f(p)}$ be a subset of $\mathcal{V}_{f(p)}(\mathscr{T}_Y)$ that generates $\mathcal{V}_{f(p)}(\mathscr{T}_Y)$. Then the continuity of f at p is equivalent to:

$$\forall U \in S_{f(p)},\ \exists V \in \mathcal{B}_p,\ f(V) \subseteq U.$$

In the case where X and Y are metric spaces, this condition becomes:

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in \mathrm{Dom}(f),\ \mathrm{d}(x,p) < \delta \Rightarrow \mathrm{d}(f(x), f(p)) < \varepsilon.$$

Proposition 8.3.3 Let $(X, \mathscr{T}_X)$, $(Y, \mathscr{T}_Y)$, $(Z, \mathscr{T}_Z)$ be topological spaces. $f : X \longrightarrow Y$, $g : Y \longrightarrow Z$ be functions. $p \in \mathrm{Dom}(g \circ f)$. Assume that f is continuous at p and g is continuous at $f(p)$. Then $g \circ f$ is continuous at p.

Proof Let U be a neighborhood of $g(f(p))$. Since g is continuous at $f(p)$, there exists a neighborhood V of $f(p)$ such that $g(V) \subseteq U$. Since f is continuous at p, there exists a neighborhood W of p, such that $f(W) \subseteq V$. Hence $g(f(W)) \subseteq g(V) \subseteq U$. So $g \circ f$ is continuous at p. □

Example 8.3.4 Let $(X, \mathscr{T}_X)$ be a topological space. Then Id_X and constant mapping are continuous.

Theorem 8.3.5 Let $(X, \mathscr{T}_X)$, $(Y, \mathscr{T}_Y)$ be topological spaces, $f : X \longrightarrow Y$ a function, $p \in \mathrm{Dom}(f)$. Consider the following conditions:
(1) f is continuous at p.
(2) For any $(x_n)_{n\in\mathbb{N}} \in \mathrm{Dom}(f)^{\mathbb{N}}$, if $\lim\limits_{n\to\infty} x_n = p$, then

$$\lim_{n\to\infty} f(x_n) = f(p).$$

One has (1)⇒(2). If p has a countable basis of neighborhoods, (2)⇒(1).

Proof Let $(x_n)_{n\in\mathbb{N}} \in \mathrm{Dom}(f)^{\mathbb{N}}$ such that $\lim\limits_{x\to\infty} x_n = p$. f is continuous at p, so for any neighborhood U of $f(p)$, there exists a neighborhood V of p, such that $f(V) \subseteq U$. Since $\lim\limits_{n\to\infty} x_n = p$, there exists $N \in \mathbb{N}$, so that for any $n \in \mathbb{N}_{>N}$, $x_n \in V$. Hence for any $n \in \mathbb{N}_{>N}$, $f(x_n) \in U$. Hence $\lim\limits_{n\to\infty} f(x_n) = f(p)$.
Assume that p has a countable basis of neighborhood. Let $(W_n)_{n\in\mathbb{N}}$ be a sequence of neighborhood of p, such that $\{W_n \mid n \in \mathbb{N}\}$ forms a neighborhood basis. For $n \in \mathbb{N}$,

$$V_n := \bigcap_{i\in\mathbb{N}_{\leq n}} W_i.$$

V_n is a neighborhood of p. If f is not continuous at p, then there exists an neighborhood of p, U, such that

$$\forall n \in \mathbb{N},\ f(V_n) \nsubseteq U.$$

We pick $x_n \in V_n$ but $f(x_n) \notin U$. For any neighborhood V of p, there exists $N \in \mathbb{N}$ such that $V_N \subseteq V$, so that $x_n \in V$ for any $n \in \mathbb{N}_{>N}$. Hence, x_n converges to p. But $f(x_n)$ cannot converge to $f(p)$. □

Lemma 8.3.6 Let $(X, \mathscr{T}_X)$ be a topological space. $V \subseteq X$. If $\forall p \in V$, V is a neighborhood of p, then $V \in \mathscr{T}_X$. In fact $\forall p \in V$, there exists $W_p \in \mathscr{T}_X$, $p \in W_p \subseteq V$. Hence

$$V = \bigcup_{p\in V} \{p\} \subseteq \bigcup_{p\in V} W_p \subseteq V.$$

Proposition 8.3.7 Let $(X, \mathscr{T}_X)$, $(Y, \mathscr{T}_Y)$ be topological spaces, $\mathcal{S} \subseteq \mathscr{T}_Y$, $\mathcal{S}$ generates $\mathscr{T}_Y$. The following statements are equivalent:
(1) f is continuous.
(2) For any $U \in \mathscr{T}_Y$, $f^{-1}(U) \in \mathscr{T}_X$.
(3) For any $U \in \mathcal{S}$, $f^{-1}(U) \in \mathscr{T}_X$.

Proof
(1)⇒(2): For any $p \in f^{-1}(U)$, one has $f(p) \in U$. Hence, there is a neighborhood V_p of p, $f(V_p) \subseteq U$, or equivalently $V_p \subseteq f^{-1}(U)$. Therefore, $f^{-1}(U) \in \mathscr{T}_X$.
(3)⇒(2):

$$\mathscr{T}'_Y = \left\{U \in \mathscr{P}(Y) \mid f^{-1}(U) \in \mathscr{T}_X\right\}$$

By definition, $\{\varnothing, Y\} \subseteq \mathscr{T}'_Y$. If $(U_1, U_2) \in \mathscr{T}'_Y \times \mathscr{T}'_Y$, then $f^{-1}(U_1 \cap U_2) = f^{-1}(U_1) \cap f^{-1}(U_2) \in \mathscr{T}_X$. So $U_1 \cap U_2 \in \mathscr{T}'_Y$. $(U_i)_{i\in I} \in (\mathscr{T}'_Y)^I$, then

$$f^{-1}\left(\bigcup_{i\in I} U_i\right) = \bigcup_{i\in I} f^{-1}(U_i) \in \mathscr{T}_X.$$

So $\mathscr{T}'_Y$ is a topology, by (3), $\mathcal{S} \subseteq \mathscr{T}'_Y \Rightarrow \mathscr{T}_Y \subseteq \mathscr{T}'_Y$. □

8.4 Initial Topology

Definition 8.4.1 Let X be a set, $((Y_i, \mathscr{T}_i))_{i\in I}$ a family of topological spaces, $(f_i : X \longrightarrow Y_i)_{i\in I}$ a family of mappings. We call **initial topology** on X induced by $(f_i)_{i\in I}$ the topology generated by

$$\bigcup_{i\in I} \left\{f_i^{-1}(U_i) \mid U_i \in \mathscr{T}_i\right\}.$$

It is the coarsest topology on X making all f_i continuous.

Proposition 8.4.2 Let $\mathscr{T}$ be the initial topology on X induced by $(f_i)_{i\in I}$.
(1)

$$\mathcal{B} = \left\{\bigcap_{j\in J} f_j^{-1}(U_j) \mid J \subseteq I \text{ finite}, \ (U_j)_{j\in J} \in \prod_{j\in J} \mathscr{T}_j\right\}$$

is topological basis that generates $\mathscr{T}$.
(2) Let $(Z, \mathscr{T}_Z)$ be topological space, $h : Z \longrightarrow X$ be a function and $p \in \mathrm{Dom}(f)$. Then h is continuous at p if and only if $\forall i \in I, \ f_i \circ h$ is continuous at p.

Proof
(1) Let

$$S = \bigcup_{i\in I} \left\{f_i^{-1}(U_i) \mid U_i \in \mathscr{T}_i\right\},$$

$\mathcal{B}'$ be the set of the intersections of all finitely elements of S (We have proved that $\mathcal{B}'$ is a basis of $\mathscr{T}$). $\mathcal{B} \subseteq \mathcal{B}'$. Let $i_1, \ldots, i_n$ elements of I, $U_{i_k} \in \mathscr{T}_{i_k}$, $J = \{i_1, \ldots, i_n\}$, $j \in J$, $A_j = \{k \in \{1, \ldots, n\} \mid i_k = j\}$, $W_j = \bigcap_{k \in A_j} U_{i_k}$.

$$\bigcap_{k=1}^{n} f_{i_k}^{-1}\left(U_{i_k}\right) = \bigcap_{j \in J} f_j^{-1}\left(W_j\right) \in \mathcal{B}.$$

(2) Since f_i is continuous at p, if h is continuous then $\forall i \in I$, $f_i \circ h$ is continuous. Assume $\forall i \in I,\ f_i \circ h$ is continuous, then

$$\forall i \in I,\ \forall U_i \in \mathscr{T}_i,\ \left(f_i \circ h\right)^{-1}\left(U_i\right) = h^{-1}\left(f_i^{-1}\left(U_i\right)\right).$$

Therefore, for any $V \in S$, $h^{-1}(V) \in \mathscr{T}_Z$. Hence h is continuous. $\square$

Example 8.4.3 Let $(X_i, \mathscr{T}_i)$ be topological spaces, $X = \prod_{i \in I} X_i$, $\pi_i : X \longrightarrow X_i$ be a projection. The initial topology on X induced by $(\pi_i)_{i \in I}$ is called the **product topology**.

8.5 Uniform Continuity

Definition 8.5.1 Let (X, d_X), (Y, d_Y) be semimetric spaces, $f : X \longrightarrow Y$ be a function, $\alpha \in \mathbb{R}_{\geq 0}$. If for any $(x_1, x_2) \in \mathrm{Dom}(f)^2$, $\mathrm{d}\left(f(x_1), f(x_2)\right) \leq \alpha \cdot \mathrm{d}(x_1, x_2)$, then we say that f is α-Lipschitzian. If there exists $\alpha \in \mathbb{R}_{\geq 0}$ such that f is α**-Lipschitzian**, then we say that f is **Lipschitzian**.
If

$$\forall \varepsilon > 0, \exists \delta > 0,\ \forall (x_1, x_2) \in \mathrm{Dom}(f)^2,\ \mathrm{d}_X(x_1, x_2) < \delta \Rightarrow \mathrm{d}(f(x_1), f(x_2)) < \varepsilon,$$

then we say that f is **uniformly continuous**.

Proposition 8.5.2 Let (X, d_X), (Y, d_Y) be semimetric spaces, $f : X \longrightarrow Y$ be a function. If f is uniformly continuous, then f continuous.

Proof Let $p \in \mathrm{Dom}(f)$. For $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\forall (x_1, x_2) \in \mathrm{Dom}(f)^2,\ \mathrm{d}_X(x_1, x_2) < \delta \Rightarrow \mathrm{d}(f(x_1), f(x_2)) < \varepsilon.$$

In particular,

$$\forall x \in \mathrm{Dom}(f),\ \exists \delta > 0,\ \mathrm{d}_X(x,p) < \delta \Rightarrow \mathrm{d}_Y(f(x), f(p)) < \varepsilon.$$

□

Definition 8.5.3 Let K be a field, we call absolute value on K any mapping,

$$|\cdot| : K \longrightarrow \mathbb{R}_{\geq 0},$$

(1) $\forall a \in K,\ a = 0_K$ if and only if $|a| = 0$.
(2) $\forall (a,b) \in K \times K,\ |ab| = |a|\,|b|$.
(3) $\forall (a,b) \in K \times K,\ |a+b| \leq |a| + |b|$.
The pair $(K, |\cdot|)$ is called a **valued field**.

Example 8.5.4 Let $(K, \leq)$ be a totally ordered field, then $|a| = \max\{-a, a\}$ is an absolute value on K.

Example 8.5.5 Let p be a prime number. Any non-zero rational number α can be written in the form

$$\alpha = p^{\mathrm{ord}_p(\alpha)} \cdot \frac{m}{n},$$

where, $\mathrm{ord}_p(\alpha) \in \mathbb{Z},\ p \nmid mn$. If $\alpha = 0$, we set (by convention) $\mathrm{ord}_p(\alpha) = +\infty$.
Properties:
(1) $\mathrm{ord}_p(\alpha\beta) = \mathrm{ord}_p(\alpha) + \mathrm{ord}_p(\beta)$.
(2) $\alpha = p^{\mathrm{ord}_p(\alpha)}\frac{m}{n},\ \beta = p^{\mathrm{ord}_p(\beta)}\frac{u}{v}$, $\mathrm{ord}_p(\alpha) > \mathrm{ord}_p(\beta)$, $p \nmid nvu$.

$$\alpha + \beta = p^{\mathrm{ord}_p(\beta)} \frac{p^{\mathrm{ord}(\alpha)-\mathrm{ord}(\beta)}mv + nu}{nv}.$$

(3) If $\mathrm{ord}(\alpha) = \mathrm{ord}(\beta)$, then $\mathrm{ord}_p(\alpha+\beta) \geq \mathrm{ord}_p(\alpha) = \mathrm{ord}_p(\beta)$.

$$\alpha + \beta = p^{\mathrm{ord}_p(\alpha)} \frac{mv + nu}{nv}.$$

Proposition 8.5.6 The mapping

$$|\cdot|_p : \mathbb{Q} \longrightarrow \mathbb{R}_{\geq 0},$$

$$\begin{cases} |\alpha|_p = p^{-\mathrm{ord}(\alpha)}, & \text{if } \alpha \neq 0 \\ |\alpha|_p = 0, & \text{if } \alpha = 0 \end{cases}$$

is an absolute value on $\mathbb{Q}$.

Proof If $\alpha = 0$, then $|\alpha|_p > 0$. If $(\alpha, \beta) \in \mathbb{Q}^2$, when $0 \in \{\alpha, \beta\}$, then $\alpha\beta = 0$ and $0 = |\alpha\beta|_p = |\alpha|_p|\beta|_p$. When $0 \notin \{\alpha, \beta\}$,

$$|\alpha\beta|_p = p^{-\mathrm{ord}_p(\alpha\beta)} = p^{-\mathrm{ord}_p(\alpha)-\mathrm{ord}_p(\beta)} = |\alpha|_p|\beta|_p.$$

If $\alpha = 0$, $|\alpha\beta|_p = |\beta|_p$. If $\beta = 0$, $|\alpha\beta|_p = |\alpha|_p$., if $0 \notin \{\alpha, \beta\}$,

$$|\alpha + \beta|_p = p^{-\mathrm{ord}_p(\alpha+\beta)} \leq p^{\max\{\mathrm{ord}_p(\alpha),\mathrm{ord}_p(\beta)\}} \leq \max\{|\alpha|_p, |\beta|_p\} \leq |\alpha|_p + |\beta|_p.$$

□

Remark 8.5.7 Let $(K, |\cdot|)$ be a valued field. If for any $(\alpha, \beta) \in K^2$ satisfies $|\alpha + \beta| \leq \max\{|\alpha|, |\beta|\}$, we say that $(K, |\cdot|)$ is **non-archimedean**, otherwise, we say that $(K, |\cdot|)$ is **archimedean**. $(\mathbb{R}, |\cdot|)$ and $(\mathbb{Q}, |\cdot|)$ are archimedean.

Definition 8.5.8 Let $(K, |\cdot|)$ be a valued filed, V a vector spaced over K. We call **seminorm** on V any mapping

$$\|\cdot\| : V \longrightarrow \mathbb{R}_{\geq 0}$$

that satisfies the following conditions:
(1) $\forall (a, x) \in K \times V,\ \|ax\| = |a| \cdot \|x\|$.
(2) $\forall (x, y) \in V \times V,\ \|x + y\| \leq \|x\| + \|y\|$.
Note that (1) implies that $\|0_V\| = |0_K| \cdot \|0_V\| = 0$.
The pair $(V, \|\cdot\|)$ is called **seminormed vector space** *over* $(K, |\cdot|)$. If $\forall (x, y) \in V \times V,\ \|x + y\| \leq \max\{\|x\|, \|y\|\}$, then we say that $\|\cdot\|$ is **ultrametric**. If $\forall x \in V\backslash\{0\}, \|x\| > 0$, then we say that $\|\cdot\|$ is a **norm** and $(V, \|\cdot\|)$ is a **normed vector space** *over* $(K, |\cdot|)$.

Example 8.5.9 $\mathrm{d} : V \times V \longrightarrow \mathbb{R}_{\geq 0}$, $\mathrm{d}(x, y) := \|x - y\|$ is a semi-metric.

Example 8.5.10 Let $(K, |\cdot|)$ be a valued field.
(1) $(K, |\cdot|)$ is a normed vector space over $(K, |\cdot|)$. ($\mathrm{d}(x, y) = |x - y|$ is a metric.)
(2) Let $(V_1, \|\cdot\|_1), \ldots, (V_n, \|\cdot\|_n)$ be seminormed vector spaces over $(K, |\cdot|)$,

$V = V_1 \oplus \cdots \oplus V_n$.

$$\|\cdot\|_{l^\infty} : V \longrightarrow \mathbb{R}_{\geq 0},\ (x_1, \ldots, x_n) \longmapsto \max_{i \in \{1,\ldots,n\}} \|x_i\|_i,\ x_i \in V_i,\ i \in \{1, \ldots, n\},$$

$$\|\cdot\|_{l^1} : V \longrightarrow \mathbb{R}_{\geq 0},\ (x_1, \ldots, x_n) \longmapsto \sum_{i \in \{1,\ldots,n\}} \|x_i\|_i,\ x_i \in V_i,\ i \in \{1, \ldots, n\}.$$

$\forall \lambda \in K,\ \forall (x_1, \ldots, x_n) \in V,$

$$\begin{aligned}
\|\lambda(x_1, \ldots, x_n)\|_{l^\infty} &= \|(\lambda \cdot x_1, \ldots, \lambda x_n)\|_{l^\infty} \\
&= \max_{i \in \{1,\ldots,n\}} \|\lambda x_i\|_i \\
&= \max_{i \in \{1,\ldots,n\}} |\lambda|\, \|x_i\|_i \\
&= |\lambda| \max_{i \in \{1,\ldots,n\}} \|x_i\|_i \\
&= |\lambda| \cdot \|(x_1, \ldots, x_n)\|_{l^\infty}.
\end{aligned}$$

$$\begin{aligned}
\|\lambda(x_1, \ldots, x_n)\|_{l^1} &= \|(\lambda \cdot x_1, \ldots, \lambda x_n)\|_{l^1} \\
&= \sum_{i \in \{1,\ldots,n\}} \|\lambda x_i\|_i \\
&= \sum_{i \in \{1,\ldots,n\}} |\lambda|\, \|x_i\|_i \\
&= |\lambda| \sum_{i \in \{1,\ldots,n\}} \|x_i\|_i \\
&= |\lambda| \cdot \|(x_1, \ldots, x_n)\|_{l^1}.
\end{aligned}$$

$\forall x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n) \in V,$

$$\begin{aligned}
\|x + y\|_{l^\infty} = \|(x_1 + y_1, \ldots, x_n + y_n)\|_{l^\infty} &= \max_{i \in \{1,\ldots,n\}} \|x_i + y_i\|_i \\
\leq \max_{i \in \{1,\ldots,n\}} \|x_i\|_i + \|y_i\|_i &\leq \|x\|_{l^\infty} + \|y\|_{l^\infty}.
\end{aligned}$$

$$\begin{aligned}
\|x + y\|_{l^1} &= \sum_{i \in \{1,\ldots,n\}} \|x_i + y_i\|_i \\
&\leq \sum_{i \in \{1,\ldots,n\}} \|x_i\|_i + \|y_i\|_i = \|x\|_{l^1} + \|y\|_{l^1}.
\end{aligned}$$

(3) Let $(V, \|\cdot\|)$ be a seminormed vector space over K, $f : W \longrightarrow V$ be a K-linear mapping. We denote by $\|\cdot\|_f$ the mapping $W \longrightarrow \mathbb{R}_{\geq 0}$ define as

$$\forall x \in W,\ \|x\|_f := \|f(x)\|.$$

$\forall(\lambda, x) \in K \times W$,

$$\|\lambda x\|_f = \|f(\lambda x)\| = \|\lambda f(x)\| = |\lambda| \cdot \|f(x)\| = \lambda \|x\|_f.$$

$\forall(x, y) \in W \times W$,

$$\|x + y\|_f = \|f(x + y)\| = \|f(x)\| + \|f(y)\| \leq \|x\|_f + \|y\|_f.$$

Therefore, $\|\cdot\|_f$ is a seminorm on W, called the **seminorm** *induced by (the K-linear) mapping f*.
(4) Let $(V, |\cdot|)$ be a seminormed vector space over K, let $\pi : V \longrightarrow E$ be a surjective K-linear mapping. We denote by $\|\cdot\|_\pi$ the mapping

$$E \longrightarrow \mathbb{R}_{\geq 0},$$

$$\alpha \longmapsto \inf_{x \in \pi^{-1}(\alpha)} \|x\|.$$

If $(\lambda, \alpha) \in K \times E$,

$$\|\lambda\alpha\|_\pi = \inf_{x \in \pi^{-1}(\lambda\alpha)} \|x\| = \inf_{x \in \pi^{-1}(\alpha)} |\lambda| \, \|x\| = |\lambda| \, \|\alpha\|_\pi.$$

If $(\alpha, \beta) \in E \times E$,

$$\begin{aligned} \|\alpha + \beta\|_\pi = \inf_{z \in \pi^{-1}(\alpha+\beta)} \|z\| &= \inf_{(x,y) \in \pi^{-1}(\alpha) \times \pi^{-1}(\beta)} \|x + y\| \\ &\leq \inf_{(x,y) \in \pi^{-1}(\alpha) \times \pi^{-1}(\beta)} \|x\| + \|y\| \\ &= \inf_{x \in \pi^{-1}(\alpha)} \|x\| + \inf_{y \in \pi^{-1}(\beta)} \|y\|. \end{aligned}$$

Hence $\|\cdot\|_\pi$ is a seminorm on E called the **quotient seminorm** *of $\|\cdot\|$ induced by π*.

Proposition 8.5.11 Let $(V, \|\cdot\|)$ be a seminormed vector space over a valued field $(K, |\cdot|)$.
(1) For any $a \in V$, the mapping $\tau_a : V \longrightarrow V$, $\tau_a(x) = x + a$ is 1-Lipschitzian.
(2) For any $\lambda \in K$, the mapping $m_\lambda : V \longrightarrow V$, $m_\lambda(x) := \lambda \cdot x$ is λ-Lipschitzian.
(3) The mapping $\|\cdot\| : V \longrightarrow \mathbb{R}$ is 1-Lipschitzian.

Proof
(1) $\forall(x, y) \in V \times V$, $\|\tau_a(x) - \tau_a(y)\| = \|(x + a) - (y + a)\| = \|x - y\|$.

(2) $\forall (x,y) \in V \times V,\ \|m_\lambda(x) - m_\lambda(y)\| = \|\lambda x - \lambda y\| = |\lambda|\,\|x-y\|$.
(3) $\forall (x,y) \in V \times V,\ \|x\| = \|(x-y)+y\| \leq \|y\| + \|x-y\|$. So $\|x\| - \|y\| \leq \|x-y\|$. Similarly, $\|y\| - \|x\| \leq \|y-x\|$. Hence,

$$|\|x\| - \|y\|| \leq \|x-y\|.$$

□

Definition 8.5.12 $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ be two seminormed vector spaces over a valued field $(K, |\cdot|)$, and φ a K-linear mapping from E to F. We define $\|\varphi\| \in [0, +\infty]$ as

$$\|\varphi\| := \sup_{\substack{x \in E \\ \|x\|_E \neq 0}} \frac{\|\varphi(x)\|_F}{\|x\|_E}.$$

In the case where $\|x\|_E = 0$, for any $x \in E$, by convention, $\|\varphi(x)\|$ is defined to be 0. If $\|\varphi\| < +\infty$, we say that φ is bounded. We denote by $\mathscr{L}(E, F)$ the set of all bounded K-linear mappings from E to F.

Remark 8.5.13 In the case when $(E, \|\cdot\|_E) = (K, \|\cdot\|)$,

$$\|\varphi\| = \sup_{x \in \mathcal{B}(0,1)} \|\varphi(x)\|_F.$$

Proposition 8.5.14
(1) For any $\varphi \in \mathscr{L}(E, F)$ be the mapping φ is $\|\varphi\|$-Lipschitzian. In particular, φ is continuous.
(2) Suppose that there exists $\lambda \in K$, such that $|\lambda| > 1$. If $\varphi : E \longrightarrow F$ is continuous at 0_E, then $\varphi \in \mathscr{L}(E, F)$.

Proof For any $(x, y) \in E \times E$:
(1) $\|\varphi(x) - \varphi(y)\|_F = \|\varphi(x-y)\|_F \leq \|\varphi\|\|x-y\|_E$.
(2) $\mathcal{B}(0_F, 1) := \{\alpha \in F \mid \|\alpha\|_F < 1\}$ is a neighborhood of 0_F. There exists $\varepsilon > 0$ such that

$$\varphi(\overline{\mathcal{B}}(0_E, \varepsilon)) \subseteq \mathcal{B}(0_F, 1)$$

where

$$\overline{\mathcal{B}}(0_E, \varepsilon) := \{x \in E \mid \|x\|_E < \varepsilon\}.$$

Let $x \in E \backslash \{0\}$, there exists $n \in \mathbb{Z}$, such that $\|\lambda^n x\|_E = |\lambda|^n \|x\|_E < \varepsilon$ and

$\|\lambda^{n+1}x\|_E = |\lambda|^{n+1}\,\|x\|_E \geq \varepsilon$. Thus,

$$\|\varphi(x)\|_F = \|\lambda^{-n}\varphi(\lambda^n x)\|_F = |\lambda|^{-n}\,\|\varphi(\lambda^n x)\|_F \leq |\lambda|^{-n} \leq \frac{|\lambda|}{\varepsilon}\|x\|_E.$$

Therefore, $\|\varphi\| \leq \frac{\lambda}{\varepsilon}$. □

Proposition 8.5.15 Let $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ be two seminormed vector spaces over a valued field $(K, |\cdot|)$. Then $\mathscr{L}(E, F)$ is a vector subspace of F^E, and $\|\cdot\|$ is a seminorm on $\mathscr{L}(E, F)$, called the operator seminorm.

Proof Let φ, ψ be to K-linear mappings from E to F. For any $x \in E$, such that $\|x\|_E \neq 0$.

$$\begin{aligned}\|(\varphi+\psi)(x)\|_F =&\|\varphi(x)+\|\psi(x)\|_F \leq \|\varphi(x)\|_F + \|\psi(x)\|_F\\ \leq&\|\varphi\|\|x\|_E + \|\psi\|\|x\|_E = (\|\varphi\| + \|\psi\|)\,\|x\|_E.\end{aligned}$$

$$\|\varphi+\psi\| \leq \|\varphi\| + \|\psi\|.$$

So, $\varphi, \psi \in \mathscr{L}(E, F) \Rightarrow \varphi + \psi \in \mathscr{L}(E, F)$. Let $\lambda \in K^\times$ and $\varphi \in \mathscr{L}(E, F)$, for any $x \in E$, $\|x\|_E \neq 0$. One has

$$\|(\lambda\varphi)(x)\|_F = \|\lambda \cdot \varphi(x)\|_F = |\lambda|\,\|\varphi(x)\|_F \leq |\lambda|\,\|\varphi\|\|x\|_E.$$

So, $\|\lambda\| \leq |\lambda|\,\|\varphi\|$, $\lambda\varphi \in \mathscr{L}(E, F)$. So $\mathscr{L}(E, F)$ is a vector subspace of F^E. Note that we can apply to λ^{-1} and $\lambda\varphi$ and get

$$\|\lambda^{-1}\cdot\lambda\varphi\| = \|\varphi\| \leq \left|\lambda^{-1}\right|\|\lambda\varphi\| = |\lambda|^{-1}\,\|\lambda\varphi\|,\ |\lambda| \leq \|\lambda\varphi\|.$$

Hence, $|\lambda|\,\|\varphi\| = \|\lambda\varphi\|$ and therefore $\|\cdot\|$ is a seminorm on $\mathscr{L}(E, F)$. □

Definition 8.5.16 Let E be a vector space over K, and $\|\cdot\|_1$ and $\|\cdot\|_2$ be seminorms on E. We say that $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent, if there exists $c_1, c_2 \in \mathbb{R}_{>0}$ such that

$$\forall x \in E,\ c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_1.$$

Proposition 8.5.17 Let E be a vector space over K, and $\|\cdot\|_1$ and $\|\cdot\|_2$ be seminorms on E. If $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent, then they define the same topology on E.

Proof Let $\mathscr{T}_1$ and $\mathscr{T}_2$ to be the topologies defined by $\|\cdot\|_1$ and $\|\cdot\|_2$, respectively. Then

$$\mathrm{Id}_E : (E, \mathscr{T}_1) \longrightarrow (E, \mathscr{T}_2)$$

is bounded. So it is continuous, so $\mathscr{T}_2 \subseteq \mathscr{T}_1$. Similarly, $\mathscr{T}_1 \subseteq \mathscr{T}_2$. $\square$

Proposition 8.5.18 Let $n \in \mathbb{N}_{\geq 1}$, and (X_i, d_i), $i \in \{1, 2, \dots, n\}$ be n seminormed vector spaces over K. Let $X = \prod_{i=1}^{n} X_i$ and

$$\mathrm{d} : X \times X \longrightarrow \mathbb{R}_{\geq 0},$$

$$\mathrm{d}\left((\mathrm{x}_1, \dots, \mathrm{x}_\mathrm{n}), (\mathrm{y}_1, \dots, \mathrm{y}_\mathrm{n})\right) \longmapsto \max_{i \in \{1, \dots, n\}} \mathrm{d}_i(x_i, y_i).$$

Then d is a semimetric on X, and the topology induced by d is the product topology of $\mathscr{T}_{\mathrm{d}_i}$ (topology induced on X_i by d_i) $i \in \{1, \dots, n\}$.

Proof
(1) $\mathrm{d}\left((x_1, \dots, x_n), (x_1, \dots, x_n)\right) = \max_{i \in \{1, \dots, n\}} \mathrm{d}_i(x_i, x_i) = 0$.
(2)

$$\begin{aligned} \mathrm{d}\left((x_1, \dots, x_n), (y_1, \dots, y_n)\right) &= \max_{i \in \{1, \dots, n\}} \mathrm{d}_i(x_i, y_i) \\ = \max_{i \in \{1, \dots, n\}} \mathrm{d}_i(y_i, x_i) &= \mathrm{d}\left((y_1, \dots, y_n), (x_1, \dots, x_n)\right). \end{aligned}$$

(3)

$$\begin{aligned} &\mathrm{d}((x_1, \dots, x_n), (z_1, \dots, z_n)) \\ =& \max_{i \in \{1, \dots, n\}} \mathrm{d}_i(x_i, z_i) \\ \leq& \max_{i \in \{1, \dots, n\}} \mathrm{d}_i(x_i, y_i) + \max_{i \in \{1, \dots, n\}} \mathrm{d}_i(y_i, z_i) \\ =& \mathrm{d}((x_1, \dots, x_n), (y_1, \dots, y_n)) + \mathrm{d}((y_1, \dots, y_n), (z_1, \dots, z_n)). \end{aligned}$$

So d is a semimetric on X. For $i \in \{1, \dots, n\}$, $\mathscr{T}_i := \mathscr{T}_{\mathrm{d}_i}$, where $\mathscr{T}_{\mathrm{d}}$ is the topology induced by d. Let $\pi_i : X \longrightarrow X_i$ be the project mapping (continuous with the product topology on X.) For any $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$ in X,

$$\mathrm{d}(x, y) = \max_{i \in \{1, \dots, n\}} \mathrm{d}_i(x_i, y_i) = \max_{i \in \{1, \dots, n\}} \mathrm{d}_i(\pi_i(x), \pi_i(y)),$$

have $\forall i \in \{1, \dots, n\}$, $\mathrm{d}_i(x_i, y_i) \leq \mathrm{d}(x, y)$ which implies that

$$\mathrm{Id}_X : (X, \mathscr{T}_{\mathrm{d}}) \longrightarrow (X, \mathscr{T})$$

is continuous. So $\mathscr{T} \subseteq \mathscr{T}_{\mathrm{d}}$.

$$\begin{aligned}\mathcal{B}((p_1, \ldots, p_n), \varepsilon) =& \{(x_1, \ldots, x_n) \mid \mathrm{d}((p_1, \ldots, p_n), (x_1, \ldots, x_n)) < \varepsilon\} \\ =& \prod_{i=1}^{n} \mathcal{B}(p_i, \varepsilon) \in \mathscr{T}.\end{aligned}$$

□

8.6 Closed Subsets

Example 8.6.1 (Review of open subsets) In $\mathbb{R}$, an interval of the form $]a, b[$ is open, since $]a, b[= \mathcal{B}(\frac{a+b}{2}, \frac{b-a}{2})$. An interval of the form $]a, +\infty[$ is open, since $]a, +\infty[= \bigcup\limits_{n \in \mathbb{N}_{\geq 1}}]a, a+n[$.

Definition 8.6.2 Let $(X, \mathscr{T})$ be a topological space. We say a subset Y of X is **closed** if $X \setminus Y$ is open.

Remark 8.6.3
(1) $\varnothing$, X are closed.
(2) If F_1, F_2 are closed subset of X, then $F_1 \cup F_2$ is closed.
(3) If $(F_i)_{i \in I}$ is a non-empty family of closed subsets of X, then $\bigcap\limits_{i \in I} F_i$ is closed.

Example 8.6.4 Let $(a, b) \in \mathbb{R}^2$, $a < b$, then $[a, b] \subseteq \mathbb{R}^2$ is closed. Moreover, $]-\infty, a]$ is closed.

Proposition 8.6.5 Let $(X, \mathscr{T}_X)$ and $(Y, \mathscr{T}_Y)$ be topological spaces, and $f : X \longrightarrow Y$ be a mapping, then the following statements are equivalent:
(1) f is continuous.
(2) For any closed subset F of Y, $f^{-1}(F)$ is a closed subset of X.

Proof
(1) $\Leftrightarrow$ (2): f is continuous if and only if, for any open subset U of Y, $f^{-1}(U) \in \mathscr{T}_X$. Let $F \subseteq Y$ be closed, then $Y \backslash F$ is open. So $f^{-1}(Y \backslash F) = X \backslash f^{-1}(F)$ is open, so $f^{-1}(F)$ is closed.
(2) $\Leftrightarrow$ (1): Let $U \in \mathscr{T}_Y$, then $F = Y \backslash U$ is closed, so $f^{-1}(F) = X \backslash f^{-1}(U)$ is

closed. So $f^{-1}(U) \in \mathscr{T}_Y$. $\square$

Example 8.6.6
In $\mathbb{R}^2$, $\{(x,y) \in \mathbb{R}^2 \mid x \geq 1\}$ is closed. Since $\mathbb{R}^2 \backslash \{(x,y) \mid x \geq 1\} =]-\infty, 0[\times \mathbb{R}$. Since $f(x,y) = x + y$ is continuous, then $f^{-1}([0,+\infty[) = \{(x,y) \in \mathbb{R}^2 \mid x + y \geq 0\}$ is closed.

Example 8.6.7 Let (X, d) be a semimetric space. Let $Y \subseteq X$, $Y \neq \varnothing$. We define, for any $x \in X$,

$$\mathrm{d}(x, Y) = \inf_{y \in Y} \mathrm{d}(x,y) \in \mathbb{R}_{\geq 0}.$$

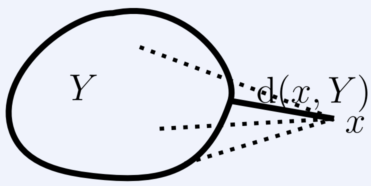


Figure 8.6.1: Definition of $\mathrm{d}(\cdot, Y)$.

The mapping $\mathrm{d}(\cdot, Y) : X \longrightarrow \mathbb{R}$, $x \longmapsto \mathrm{d}(x, Y)$ is 1-Lipschitzian.
Let $(x, x') \in X \times X$, $\forall y \in Y$,

$$\mathrm{d}(x, Y) - \mathrm{d}(x', y) \leq \mathrm{d}(x, y) - \mathrm{d}(x', y) \leq \mathrm{d}(x, x').$$

Taking the supremum, we get

$$\mathrm{d}(x, Y) - \mathrm{d}(x', Y) \leq \mathrm{d}(x, x').$$

By symmetry between x and x', $\mathrm{d}(x', Y) - \mathrm{d}(x, Y) \leq \mathrm{d}(x, x')$. So

$$|\mathrm{d}(x, Y) - \mathrm{d}(x', Y)| \leq |\mathrm{d}(x, x')|.$$

For any $r > 0$, define

$$B(Y, r) := \{x \in X \mid \mathrm{d}(x, Y) < r\},$$

$$\overline{B}(Y, r) := \{x \in X \mid \mathrm{d}(x, Y) \leq r\}$$

$B(Y, r)$ is open, $\overline{B}(Y, r)$ is closed. If $Y = \{y\}$ is a one point set, $B(Y, r)$ and $\overline{B}(Y, r)$ are defined as $B(y, r)$ and $\overline{B}(y, r)$ respectively.

Definition 8.6.8 Let $(X, \mathscr{T})$ be a topological space.
(1) Let $\mathcal{F}$ be a non-degenerate filter, an element $p \in X$ is called **adherent point** *of* $\mathcal{F}$ if $\mathcal{F} \cup \mathcal{V}_p(\mathscr{T})$ generates a non-degenerate filter. ($\forall U \in \mathcal{F}, \forall V \in \mathcal{V}_p(\mathscr{T}), U \cap V \neq \varnothing$.)
(2) Let $Y \subseteq X$. We say that $p \in X$ is an adherent point of Y if it is an adherent point of the principal filter $\mathcal{F}_Y = \{U \subseteq X \mid Y \subseteq U\}$. (For any neighborhood V of p, $Y \cap V \neq \varnothing$.) We denote by $\overline{Y}$ the set of all adherent points of Y called the **closure** of Y. Clearly, $Y \subseteq \overline{Y}$.

Proposition 8.6.9 Let $(X, \mathscr{T}_X)$ be a topological space, $Y \subseteq X$. Then $\overline{Y}$ is the smallest closed subset containing Y. Namely,

$$\overline{Y} = \bigcap_{\substack{Y \subseteq F \\ F \subseteq X \text{ closed},}} F.$$

Proof Let $p \in \overline{Y}$. If there exists a closed subset F containing Y such that $p \notin F$. So $p \in X \backslash F$. Hence $X \backslash F \in \mathcal{V}_p(\mathscr{T})$. So $\varnothing = (X \backslash Y) \cap Y \supseteq (X \backslash F) \cap Y \neq \varnothing$. Contradiction. Therefore,

$$\overline{Y} \subseteq \bigcap_{\substack{Y \subseteq F \\ F \subseteq X \text{ closed},}} F.$$

Suppose that $x \in X \backslash \overline{Y}$. There exists an open neighborhood U of x such that $U \cap Y = \varnothing$. So $x \notin F := X \backslash U$.

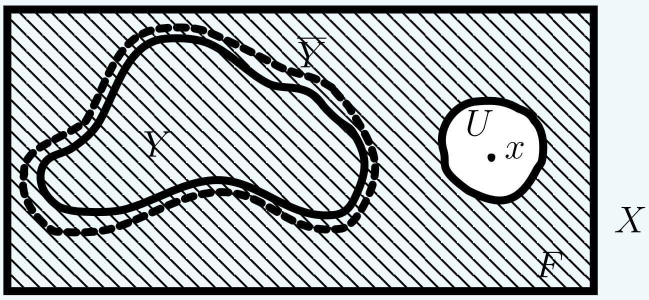


Figure 8.6.2: Closure

Note that F is closed and $F \supseteq Y$. Therefore,

$$X \backslash \overline{Y} \subseteq \bigcup_{\substack{Y \subseteq F \\ F \subseteq X \text{ closed},}} (X \backslash F).$$

which leads to

$$\overline{Y} \supseteq X \backslash \left(\bigcup_{\substack{Y \subseteq F \\ F \subseteq X \text{ closed,}}} (X \backslash F) \right) = \bigcap_{\substack{Y \subseteq F \\ F \subseteq X \text{ closed,}}} F.$$

□

Definition 8.6.10 Let $(X, \mathscr{T})$ be a topological space and $Y \subseteq X$. We denote by Y° the set of $p \in Y$ such that Y is a neighborhood of p.

Proposition 8.6.11 Y° is the least[a] open subset of X such that is contained in Y. Moreover,

$$X \backslash Y^\circ = \overline{X \backslash Y}.$$

[a]largest

Proof $\forall y \in Y^\circ$, there exists $U_y \in \mathscr{T}$ such that $y \in U_y \subseteq Y$. Therefore, $\forall x \in U_y$, Y is a neighborhood of x, hence, $U_y \subseteq Y^\circ$. We thus obtain

$$Y^\circ = \bigcup_{y \in Y^\circ} \{y\} \subseteq \bigcup_{y \in Y^\circ} U_y \subseteq Y^\circ.$$

Hence Y° is open.
If $U \subseteq Y$ is open, then $\forall x \in U$, Y is a neighborhood of x. So $U \subseteq Y^\circ$. Therefore, Y° is the largest open subset that is contained in Y.

$$X \backslash Y^\circ = X \backslash \bigcup_{\substack{U \subseteq Y \\ U \in \mathscr{T}}} U = \bigcap_{\substack{U \subseteq Y \\ \in \mathscr{T}}} X \backslash U \overset{F = X \backslash U}{=} \bigcap_{\substack{F \text{ closed} \\ X \backslash Y \subseteq F}} F = \overline{X \backslash Y}.$$

□

Definition 8.6.12 Let $(X, \mathscr{T})$ be a topological space. We equip $X \times X$ with the product topology, (a topological basis is given by $\{U \times V \mid (U, V) \in \mathscr{T}^2\}$) Let

$$\Delta_X := \{(x, x) \mid x \in X\} \subseteq X \times X.$$

If Δ_X is closed, we say that $(X, \mathscr{T})$ is a **Hausdorff space**. (Or $(X, \mathscr{T})$ is separated.)

Proposition 8.6.13 $(X, \mathscr{T})$ is a Hausdorff space if and only if $\forall (x, y) \in X \times X$, $x \neq y$, there exists $(U, V) \in \mathcal{V}_x(\mathscr{T}) \times \mathcal{V}_y(\mathscr{T})$, such that $U \cap V = \varnothing$.

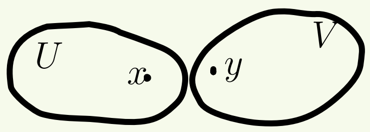


Figure 8.6.3: Hausdorff space

Proof
"$\Rightarrow$": If $(x, y) \in X \times X$, $x \neq y$, then $(x, y) \in (X \times X) \backslash \Delta_X$. There exists $(U, V) \subseteq (X \times X) \backslash \Delta_X$, such that $(x, y) \in U \times V$, so $(U \times V) \cap \Delta_X = \varnothing$. Thus $U \cap V = \varnothing$. (If $p \in U \cap V$, then $(p, p) \in (U \times V) \cap \Delta_X$.)
"$\Leftarrow$": For any $(x, y) \in X \times X$, $x \neq y$, $\exists U \in \mathscr{T}, \forall V \in \mathscr{T}, x \in U, y \in V, U \cap V = \varnothing$. Then $(x, y) \in U \times V$ and $(U \times V) \cap \Delta_X = \varnothing$. So Δ_X is closed. $\square$

Proposition 8.6.14 Let $(X, \mathscr{T})$ be a Hausdorff space. Let $\mathcal{F}$ be a non-degenerate filter on X. If $\mathcal{F}$ has a limit point[a], then its limit point is unique.

[a] Let $(X, \mathscr{T})$ be a topological space and $\mathcal{F}$ be a filter on X. If $p \in X$ is such that $\mathcal{V}_p(\mathscr{T}) \subseteq \mathcal{F}$, then we say that p is a limit point of $\mathcal{F}$.

Proof (By contradiction) Suppose that x and y are limit points of $\mathcal{F}$, $x \neq y$. Since X is Hausdorff, $\exists (U, V) \in \mathscr{T}^2$, $x \in U, y \in V, U \cap V = \varnothing$. Since x and y are limit points of $\mathcal{F}$, $U \in \mathcal{F}, V \in \mathcal{F}$. This contradicts the hypothesis that $\mathcal{F}$ is non-degenerate. $\square$

Example 8.6.15 Any metric space is Hausdorff.
Let (X, d) be a metric space, $\forall (x, y) \in X \times X, x \neq y, \mathrm{d}(x, y) > 0$. Let $\varepsilon = \frac{\mathrm{d}(x,y)}{2}$. $B(x, \varepsilon) \cap B(y, \varepsilon) = \varnothing$. In fact, if $z \in B(x, \varepsilon) \cap B(y, \varepsilon)$,

$$\mathrm{d}(x, y) \leq \mathrm{d}(x, z) + \mathrm{d}(z, y) < 2\varepsilon = \mathrm{d}(x, y).$$

Proposition 8.6.16 Let $(X, \mathscr{T})$ be a topological space, Y be a subset of X and $p \in X$.
(1) Let Z be a set and $f : Z \longrightarrow X$ be a mapping. Let $\mathcal{F}$ be a non-degenerate filter on Z. If p is a limit of f along $\mathcal{F}$, and if $f(Z) \subseteq Y$, then $p \in \overline{Y}$.
(2) Suppose that p has a countable neighborhood basis. If $p \in \overline{Y}$, then there exists a sequence $(y_n)_{n \in \mathbb{N}}$ in Y that converges to p.

Proof
(1) p is a limit of f along $\mathcal{F}$ if and only if $\mathcal{V}_p(\mathscr{T}) \subseteq f_*(\mathcal{F})$, or equivalently

$$\forall U \in \mathcal{V}_p(\mathscr{T}), f^{-1}(U) \in \mathcal{F}.$$

$f(f^{-1}(U)) \subseteq U \cap Y$, since $f(X) \subseteq Y$. Hence $U \cap Y \neq \varnothing$. So $p \in \overline{Y}$.
(2) Since p has a countable neighborhood basis, there exists a decreasing sequence $V_0 \supseteq V_1 \supseteq \dots$ of neighborhood of p such that $\{V_n \mid n \in \mathbb{N}\}$ forms a filter basis of $\mathcal{V}_p(\mathscr{T})$. For any $n \in \mathbb{N}$, $V_n \cap Y = \varnothing$, we take $y_n \in V_n \cap Y$. The sequence $(y_n)_{n\in\mathbb{N}}$ converges to p since $\forall n \in \mathbb{N}$, $\{y_k \mid k \in \mathbb{N}, k \geq n\} \subseteq V_n$. □

Example 8.6.17 Let (X, d) be a semimetric space. $Y \subseteq X$, $\varepsilon > 0$. If $(y_n)_{n\in\mathbb{N}}$ is a sequence in $B(Y, \varepsilon)$, that converges to some $p \in X$, then

$$\lim_{n\to\infty} \mathrm{d}(y_n, Y) = \mathrm{d}(p, Y).$$

Therefore, $\overline{B(Y, \varepsilon)} \subseteq \overline{B}(Y, \varepsilon) := \{x \in X \mid \mathrm{d}(x, Y) \leq \varepsilon\}$.

Proposition 8.6.18 Let (X, d) be a semimetric space, $Y \subseteq X$ be a closed subset. $\forall x \in X \backslash Y$, $\mathrm{d}(x, Y) > 0$.

Proof $X \backslash Y$ is open, so $\exists \varepsilon > 0$ such that $B(X, \varepsilon) \subseteq X \backslash Y$. So $\forall y \in Y$, $\mathrm{d}(x, y) \geq \varepsilon$. Hence, $\mathrm{d}(x, y) \geq \varepsilon$. □

Corollary 8.6.19 Let $(V, \|\cdot\|)$ be a semimetric space, W be a closed vector subspace of V. $Q = V/W$. Then the quotient seminorm

$$\|\cdot\|_Q : Q \longrightarrow R,$$

$$\alpha \longmapsto \inf_{x\in V, [x]=\alpha} \|x\|$$

is a norm.

Proof Let $\alpha \in Q\backslash\alpha$ and $x \in V$ such that $\alpha = [x]$. Since $\alpha \neq 0, x \notin W$.

$$0 < \mathrm{d}(x, W) := \inf_{y\in W} \|x - y\| = \inf_{\substack{x'\in V \\ [x']=\alpha}} \|x'\| = \|\alpha\|_Q.$$

□

Proposition 8.6.20 If $(X, \mathscr{T})$ is a Hausdorff space, then, $\forall x \in X$, $\{x\}$ is closed.

Proof $\forall y \in X \backslash \{x\}, y \neq x$. So $\exists (U, V) \in \mathscr{T} \times \mathscr{T}, x \in U, y \in V$. $U \cap V = \varnothing$. So $V \subseteq X \backslash \{x\}$. Hence $X \backslash \{x\}$ is a neighborhood of y. □

Remark 8.6.21 Let $(V, \|\cdot\|)$ be a seminorm space. $W \subseteq V, Q = V/W$ and $\|\cdot\|_Q$ is the quotient seminorm. The mapping $\pi : V \longrightarrow Q, x \longmapsto [x]$ is continuous since $\|[x]\|_Q \leq \|x\|$. If $\|\cdot\|_Q$ is a norm then $\{0_Q\}$ is closed (since Q is Hausdorff). So $W = \pi^{-1}(\{0_Q\}) = \ker(\pi)$ is closed. This shows that $\|\cdot\|_Q$ is a norm $\Leftrightarrow W$ is closed.

8.7 Completeness

Definition 8.7.1 Let (X, d) be a semimetric space, $Y \subseteq X$, we define the diameter of Y as $\mathrm{diam}(Y) := \sup_{(x,y) \in Y^2} \mathrm{d}(x, y)$. If $\mathrm{diam}(Y) < +\infty$, we say that Y is **bounded**.

Remark 8.7.2 Let (E, d) be a semimetric space.
(1) If A and B are subsets of E, then

$$A \subseteq B \Rightarrow \mathrm{diam}(A) \leq \mathrm{diam}(B).$$

(2) If $A \subseteq B \subseteq E$ and B is bounded, then A is bounded.
(3) If $A = \{x_1, \ldots, x_n\} \subseteq E$, $n \in \mathbb{N}_{\geq 1}$. Then

$$\mathrm{diam}(A) = \max_{(i,j) \in \{1,\ldots,n\}^2} \mathrm{d}(x_i, x_j) < +\infty.$$

So A is bounded.
(4) $\forall p \in X$, $\mathrm{diam}(\bar{B}(p, r)) \leq 2r$, $\forall r \in \mathbb{R}_{>0}$. In fact, $\forall (x, y) \in \bar{B}(p, r)^2$, $\mathrm{d}(x, y) \leq \mathrm{d}(x, p) + \mathrm{d}(p, y) \leq 2r$.

Proposition 8.7.3 Let (E, d) be a semimetric space, $A \subseteq E$. Suppose that A is bounded. Let $r = \mathrm{diam}(A)$. For any $p \in A$, $A \subseteq \bar{B}(p, r)$.

Proof $\forall x \in A$, $\mathrm{d}(p, x) \leq \mathrm{diam}(A) = r$. □

Proposition 8.7.4 Let (E, d) be a semimetric space, $A \subseteq E$, $B \subseteq E$ and $(x_0, y_0) \in A \times B$. Then

$$\operatorname{diam}(A \cup B) \leq \operatorname{diam}(A) + \operatorname{diam}(B) + \mathrm{d}(x_0, y_0).$$

Proof Let $(x, y) \in (A \cup B)^2$.
Case 1. $\{x, y\} \in A$, $\mathrm{d}(x, y) \leq \operatorname{diam}(A)$.
Case 2. $\{x, y\} \in B$, $\mathrm{d}(x, y) \leq \operatorname{diam}(B)$.
Case 3. $x \in A,\ y \in B$,

$$\mathrm{d}(x, y) \leq \mathrm{d}(x, x_0) + \mathrm{d}(x_0, y_0) + \mathrm{d}(y_0, y) \leq \operatorname{diam}(A) + \operatorname{diam}(B) + \mathrm{d}(x_0, y_0).$$

Case 4. $x \in B,\ y \in A$. Same as case 3.

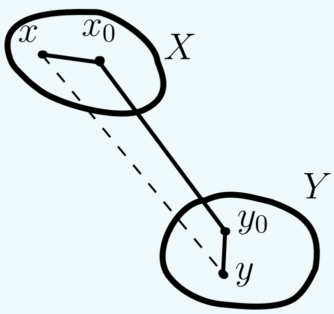


Figure 8.7.1: proposition. 8.7.4

□

Corollary 8.7.5 If $A \cap B \neq \varnothing$, then

$$\operatorname{diam}(A \cup B) \leq \operatorname{diam}(A) + \operatorname{diam}(B).$$

Definition 8.7.6 Let (E, d) be a semimetric space, and $\mathcal{F}$ be a non-degenerate filter on E. If $\inf_{A \in \mathcal{F}} \operatorname{diam}(A) = 0$, we say that $\mathcal{F}$ is a **Cauchy filter**.

Example 8.7.7 $E =]0, 1]$. If $\mathcal{F}$ is generated by $]0, \varepsilon[$, $\varepsilon \leq 1$, then $\mathcal{F}$ is a Cauchy filter.

Proposition 8.7.8 Let (E, d) be a semimetric space and $\mathcal{F}$ be a non-degenerate filter on E. If $\mathcal{F}$ has a limit point p, then $\mathcal{F}$ is a Cauchy filter.

Proof $\forall \varepsilon > 0$, $\bar{B}(p, \frac{\varepsilon}{2}) \in \mathcal{F}$, $\operatorname{diam}(\bar{B}(p, \frac{\varepsilon}{2})) \leq \varepsilon$. So $\inf\limits_{A \in \mathcal{F}} \operatorname{diam}(A) = 0$. □

Proposition 8.7.9 Let (E, d) be a semimetric space, and $\mathcal{F}$ be a Cauchy filter on E. Any adherent point of $\mathcal{F}$ is a limit point of $\mathcal{F}$.

Proof Let $\varepsilon > 0$. Let $A \in \mathcal{F}$ such that $\operatorname{diam}(A) < \frac{\varepsilon}{2}$. Note that $B(p, \frac{\varepsilon}{4}) \cap A \neq \varnothing$. So

$$\operatorname{diam}(B(p, \frac{\varepsilon}{4}) \cup A) \leq \operatorname{diam}(B(p, \frac{\varepsilon}{4})) + \operatorname{diam}(A) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.$$

So $A \subseteq B(p, \varepsilon)$. □

Definition 8.7.10 Let (E, d) be a semimetric space. $I \subseteq \mathbb{N}$ be an infinite subset, $\mathcal{F}$ be the Fréchet filter on I. We say that a sequence $x := (x_n)_{n \in I} \in E^I$ is a Cauchy sequence if $x_*(\mathcal{F}) := \{U \subseteq E \mid x^{-1}(U) \in \mathcal{F}\}$ is a Cauchy filter. Or equivalently,

$$\forall \varepsilon > 0, \ \exists N \in I, \ \forall n, m \in I_{\geq N}^2, \ \mathrm{d}(x_n, x_m) \leq \varepsilon.$$

Proof "⇒" $\forall \varepsilon > 0, \ \exists U \in x_*(\mathcal{F}), \ \operatorname{diam}(U) \leq \varepsilon$. Since $x^{-1}(U) \in \mathcal{F}$, $\exists N \in \mathbb{N}$ such that $I_{\geq N} \subseteq x^{-1}(U)$. So $\{x_n \mid n \in I_{\geq N}\} \subseteq U$, which leads to

$$\forall (n, m) \in I_{\geq N}^2, \ \mathrm{d}(x_n, x_m) \leq \varepsilon.$$

"⇐" $\forall \varepsilon > 0, \ \exists N \in \mathbb{N}$, $\operatorname{diam}\left(\{x_n \mid n \in I_{\geq N}\}\right) \leq \varepsilon$. □

Proposition 8.7.11 Let (E, d) be a semimetric space, and $(x_n)_{n \in I}$ be a sequence in E.
(1) If $(x_n)_{n \in I}$ is convergent, then it is a Cauchy sequence.
(2) If $(x_n)_{n \in I}$ is a Cauchy sequence, then $\{x_n \mid n \in I\}$ is bounded.
(3) If $(x_n)_{n \in I}$ is a Cauchy sequence, then any of its subsequence is a Cauchy sequence.
(4) If $(x_n)_{n \in I}$ is a Cauchy sequence, and if there exists a subsequence $(x_n)_{n \in J}$ that converges to some $p \in E$, then $(x_n)_{n \in I}$ converges to p.

Proof Let $\mathcal{F}_I$ be a Fréchet filter on I.
(1) $x_*(\mathcal{F}_I)$ has a limit point. So it is a Cauchy filter.

(2) $\exists N \in \mathbb{N}$ such that $\operatorname{diam}(\{x_n \mid n \in I_{\geq \mathbb{N}}\}) \leq 1 < +\infty$. So $\operatorname{diam}(\{x_n \mid n \in I\}) < +\infty$ since

$$\{x_n \mid n \in I\} = \{x_n \mid n \in I_{<N}\} \cup \{x_n \mid n \in I_{\geq N}\}.$$

(3) Let J be an infinite subset of I, $\lambda : J \longrightarrow I$ be the inclusion mapping, and $\mathcal{F}_J$ be the Fréchet filter. Then $\lambda_*(\mathcal{F}_J) \supseteq \mathcal{F}_I$.

$$(x \circ \lambda)_*(\mathcal{F}_J) = x_*(\lambda_*(\mathcal{F}_J)) \supseteq x_*(\mathcal{F}_I.)$$

Since $x_*(\mathcal{F}_I)$ is a Cauchy filter, $(x \circ \lambda)_*(\mathcal{F}_J)$ is a Cauchy filter.
(4) We keep the notation introduced in the proof of (3). If p is a limit point of $x \circ \lambda_*(\mathcal{F}_J)$. Then p is an adherent point of $x_*(\mathcal{F}_I)$. Since $x_*(\mathcal{F}_I)$ is a Cauchy filter, p is a limit point of $x_*(\mathcal{F}_I)$. □

Definition 8.7.12 Let (X, d) be a semimetric space. If any Cauchy filter on X has a limit point, then we say that (X, d) is complete.

Proposition 8.7.13 Let (X, d) be a metric space and Y be a subset of X. If (Y, d) is complete, then Y is a closed subset of X.

Proof Let $(y_n)_{n\in\mathbb{N}}$ be a sequence in Y that converges in X to some $p \in X$. So $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in Y, thus it converges to some $q \in Y$. Since X is Hausdorff, $p = q$. Then $\bar{Y} = Y$ (Since $\bar{Y}$ is the set of limits of all sequence in Y). So Y is closed. □

Example 8.7.14 $(\mathbb{R}, |\cdot|)$ is complete.
Let $(x_n)_{n\in I}$ be a Cauchy sequence in $\mathbb{R}$. Let $M > 0$ such that $\forall n \in I, |x_n| \leq M$. Hence,

$$-M \leq \liminf_{n\to+\infty} x_n \leq \limsup_{n\to+\infty} x_n \leq M.$$

By Bolzano-Weierstrass, $\exists J \subseteq I$ infinite such that $(x_n)_{n\in I}$ converges to $\limsup_{n\to+\infty} x_n \in \mathbb{R}$. So $(x_n)_{n\in I}$ converges.

Proposition 8.7.15 Let (X, d) be a semimetric space, (X, d) is complete if and only if any Cauchy sequence in X is convergent.

Proof It is suffices to prove "$\Leftarrow$". Suppose that all Cauchy sequence in X converges. Let $\mathcal{F}$ be a Cauchy filter on X. $\forall n \in \mathbb{N}$, let $A_n \in \mathcal{F}$, $\operatorname{diam}(A) < \frac{1}{n+1}$. $\forall n \in \mathbb{N}$, let $B_n = A_0 \cap \ldots \cap A_n \in \mathcal{F}$, $\operatorname{diam}(B) \leq \frac{1}{n+1}$. Take $x_n \in B_n$. Then $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Hence $(x_n)_{n\in\mathbb{N}}$ converges to $p \in X$.

We claim that p is a limit point of $\mathcal{F}$. If V is a neighborhood of p, then $\exists N \in \mathbb{N}$ such that $\{x_n \mid n \in \mathbb{N}_{\geq N}\} \subseteq V$. So $V \cap B_n \neq \varnothing$, $\forall n \in \mathbb{N}_{\geq N}$.

Let $A \in \mathcal{F}$, $A \cap B_n \neq \varnothing$. Take $y_n \in A \cap B_n$, $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. $\mathrm{d}(y_n, x_n) \leq \frac{1}{1+n}$, so $(y_n)_{n\in\mathbb{N}}$ converges to p. Thus $V \cap A \neq \varnothing$. $\square$

Proposition 8.7.16 Let n be a positive integer. For any $i \in \{1, \ldots, n\}$, let (X_i, d_i) be a semimetric space. Let $X = X_1 \times \ldots \times X_n$ and

$$\pi_i : X \longrightarrow X_i$$

$$(x_1, \ldots, x_n) \longmapsto x_i$$

be the projection mapping. We equip X with the product semimetric d

$$\mathrm{d}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := \max_{i\in\{1,\ldots,n\}} \mathrm{d}_i(x_i, y_i).$$

Let $\mathcal{F}$ be a non-degenerate filter on X. For any $i \in \{1, \ldots, n\}$, let

$$\mathcal{F}_i = (\pi_i)_* (\mathcal{F}).$$

(1) The filter $\mathcal{F}$ is a Cauchy filter if and only if $\mathcal{F}_i$ is a Cauchy filter for any $i \in \{1, \ldots, n\}$.
(2) The filter $\mathcal{F}$ has a limit point if and only if $\mathcal{F}_i$ has a limit point for any $i \in \{1, \ldots, n\}$. If p_i is the limit point of $\mathcal{F}_i$, then $p = (p_1, \ldots, p_n)$ is a limit point of $\mathcal{F}$.
(3) If $(X_1, \mathrm{d}_1), \ldots, (X_n, \mathrm{d}_n)$ are complete, then (X, d) is complete.

Proof
(1) Since π_i are Lipschitzian, if $\mathcal{F}$ is a Cauchy filter, then $\mathcal{F}_i$ are all Cauchy filters. Conversely, let $\mathcal{F}$ be a non-degenerate filter such that $\mathcal{F}_i$ is a Cauchy filter for all $i \in \{1, ., n\}$. For any $\varepsilon > 0$, any $i \in \{1, \ldots, n\}$, there exists $A_i \in \mathcal{F}_i$ such that $\operatorname{diam}(A_i) < \varepsilon$. We define

$$A := A_1 \times \ldots \times A_n = \bigcap_{i=1}^{n} \pi_i^{-1}(A_i) \in \mathcal{F}.$$

For any $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in A$, $\mathrm{d}(x,y) = \max_{i \in \{1,\ldots,n\}} \mathrm{d}_i(x_i, y_i) < \varepsilon$.
(2) π_i is Lipschitzian so also continuous $i \in \{1, \ldots, n\}$, therefore, if $p = (p_1, \ldots, p_n)$ a limit point of $\mathcal{F}$, then $\pi_i(p) = p_i$ is a limit point of $\mathcal{F}_i$.

Conversely, Suppose that p_i is a limit point of $\mathcal{F}_i$. Let U be a neighborhood of p. There exists $U_i \in \mathcal{F}_i$ neighborhood of p_i such that $U_1 \times \ldots \times U_n \subseteq U$.

$$U \supseteq \bigcap_{i=1}^{n} \pi_i^{-1}(U_i) \in \mathcal{F},$$

so $U \in \mathcal{F}$. □

Proposition 8.7.17 Let (X, d_X), (Y, d_Y) be two semimetric spaces, $f : X \longrightarrow Y$ be uniformly continuous. For any Cauchy filter $\mathcal{F}$ on X, $f_*(\mathcal{F})$ is also a Cauchy filter on Y.

Proof $\forall \varepsilon > 0,\ \exists \delta > 0$ such that

$$\forall (x,y) \in X \times X,\ \mathrm{d}_X(x,y) < \delta \Rightarrow \mathrm{d}_Y(f(x), f(y)) < \varepsilon.$$

Let $A \in \mathcal{F}$ such that $\mathrm{diam}(A) < \delta$, then $f(A) \in f_*(\mathcal{F})$ and $\mathrm{diam}(f(A)) < \varepsilon$. □

Proposition 8.7.18 Let $(X, \mathrm{d}), (X', \mathrm{d}')$ be two semimetric spaces, $f : X \longrightarrow X'$ injective mapping. Assume that there exists positive constants C_1 and C_2 such that $\forall (x,y) \in X \times X$, $C_1 \cdot \mathrm{d}(x,y) \leq \mathrm{d}'(f(x), f(y)) \leq C_2 \cdot \mathrm{d}(x,y)$.
(1) For any sequence $(x_i)_{i \in I}$ in X, the sequence $(x_i)_{i \in I}$ is a Cauchy sequence if and only if the sequence $(f(x_i))_{i \in I}$ is a Cauchy sequence.
(2) For any $(x_i)_{i \in I}$ in X, $(x_i)_{i \in I}$ is a convergent if and only if $(f(x_i))_{i \in I}$ is convergent.
(3) If (X, d) is complete, then $(f(X), \mathrm{d}')$ is complete. $f(X)$ is closed in X' if in addition we assume that d' is a metric.

Proof Note that $f : X \longrightarrow f(X)$, $f^{-1} : f(X) \longrightarrow X$ are Lipschitzian. Apply the previous proposition. □

Theorem 8.7.19 Let $(K, |\cdot|)$ be a valued filed such that $(K, |\cdot|)$ is complete. Let V be a finite dimensional vector space over K.
(1) All possible norms on K are equivalent.
(2) For any norm $\|\cdot\|$ on V, the normed space $(V, \|\cdot\|)$ is complete.

(3) If we equip V with the topology induced by an arbitrary norm $\|\cdot\|$. For any normed vector space $(V, \|\cdot\|)$, any K-linear mapping $f : V \longrightarrow V'$ is bounded and $f(V)$ is closed in V'.

Proof Let $e := (e_i)_{i=1}^n$ be a basis of V. Then, the mapping

$$\|\cdot\|_e : V \longrightarrow \mathbb{R}_{>0},\ \|a_1e_1 + \ldots + a_ne_n\| = \max_{i\in\{1,\ldots,n\}} |e_i|$$

is a norm on V. For $f : V \longrightarrow V'$,

$$\begin{aligned} \|f(a_1e_1 + \ldots + a_ne_n)\|' &= \|a_1f(e_1) + \ldots + a_nf(e_n)\|' && (8.7.1)\\ &\le |a_1|\,\|f(e_1)\|' + \ldots + |a_n|\,\|f(e_n)\|' && (8.7.2)\\ &\le \max_{i\in\{1,\ldots,n\}} |a_i| \sum_{i=1}^n \|f(e_i)\|'. && (8.7.3) \end{aligned}$$

Therefore the K-linear mapping $f : (V, \|\cdot\|_e) \longrightarrow (V', \|\cdot\|')$ is bounded. In particular, for any $\|\cdot\|$ on V, $\mathrm{Id} : (V, \|\cdot\|) \longrightarrow (V, \|\cdot\|)$ is bounded. So there exists $C > 0$, $\|\cdot\| \le C\|\cdot\|_e$.

To prove the theorem, we reason by induction with respect to $n = \dim(V)$.

In case when $n = 0$, $V = \{0\}$ has the unique norm $\|\cdot\|_0$ (constant mapping with the value $0 \in \mathbb{R}$). Any sequence in $\{0_V\}$ is constant, and hence convergent ($(\{0\}, \|\cdot\|_0)$ complete). If $f : (\{0_V\}, \|\cdot\|) \longrightarrow (V, \|\cdot\|)$ is K-linear, then it is bounded. The unique $f(\{0_V\})$ is a one-point set, which is closed, since $(V, \|\cdot\|)$ is Hausdorff.

In case when $n = 1$, let e_1 be the basis of V, and $\|\cdot\|$ be an arbitrary norm on V.

$$\|a_1e_1\| = |a_1|\,\|e_1\| = \|a_1e_1\|_{e_1} \cdot \|e_1\|,$$

so $\|\cdot\|$ is equivalent to $\|\cdot\|_{e_1}$.
(2) Since $(K, |\cdot|)$ is complete, so $(V, \|\cdot\|_{e_1})$ is also complete, since

$$(K, |\cdot|) \longrightarrow (V, \|\cdot\|_{e_1})$$

$$a \longmapsto ae_1$$

is an isomorphism.
(3) We have seen that the mapping $f : (V, \|\cdot\|_{e_1}) \longrightarrow (V', \|\cdot\|')$ is bounded. So $f : (V, \|\cdot\|) \longrightarrow (V', \|\cdot\|')$ is also bounded. $f(V) = \{0\}$, or $(f(V), \|\cdot\|')$ is of dimension 1, so $(f(V), \|\cdot\|')$ is complete, so $f(V)$ is closed.
Suppose that the theorem holds for normed vector space of dimension $< n$. Then the case of dimension n:

Let $e = (e_i)_{i=1}^n$ be a basis of V. Let $W = \text{span}_K(\{e_1, \ldots, e_n\})$. If $\|\cdot\|$ is a norm on V, then, by the induction hypothesis,
(i) $\|\cdot\|$ and $\|\cdot\|_{\underline{e}}$ are equivalent on W, that is, $\exists A > 0$ such that $\forall (a_1, \ldots, a_{n-1})$,

$$\max\{|a_1|, \ldots, |a_n|\} = \|a_1e_1 + \ldots + a_{n-1}e_{n-1}\|_{\underline{e}} \leq A\|a_1e_1 + \ldots + a_{n-1}e_{n-1}\|.$$

(ii) $(W, \|\cdot\|)$ is complete.
(iii) W is a closed subset of V.
Let $Q = V/W$, and $\|\cdot\|_Q$ be the quotient norm on Q. Let $(b_1, \ldots, b_n) \in K^n$,

$$s = b_1e_1 + \ldots + b_ne_n \in V,\ t = b_1e_1 + \ldots + b_ne_n + w \in V,\ \alpha = [s] = b_n[e_n] \in Q.$$

$$\|t\| = \|s - b_ne_n\| \leq \|s\| + |b_n|\,\|e_n\|.$$

Take $B = \frac{\|e_n\|}{\|[e_n]\|_Q} \in \mathbb{R}$, $\|s\| \geq \|\alpha\|_Q = |b_n|\,\|[e_n]\|_Q = B^{-1}|b_n|\|e_n\|$.

$$B^{-1}\|s\| \geq (\|t\| - |b_n| \cdot \|e_n\|)\, B^{-1}$$

$$\|s\| \geq B^{-1} \cdot |b_n| \cdot \|e_n\|.$$

$$\left(B^{-1} + 1\right)\|s\| \geq B^{-1}\|t\| \geq B^{-1}A^{-1}\max\{|b_1|, \ldots, |b_n|\}$$

Take $C = \min\left\{\frac{B^{-1}A^{-1}}{B^{-1}+1}, B^{-1} \cdot \|e_n\|\right\}$. Then $\|s\| \geq C\max\{|b_1|, \ldots, |b_n|\}$. So $\|\cdot\|$ is equivalent to $\|\cdot\|_{\underline{e}}$.
(2) Since $(V, \|\cdot\|)_e$ is complete and $\|\cdot\|$ is equivalent to $\|\cdot\|_e$, we obtain that $(V, \|\cdot\|)$ is also complete.
(3) Since $f : (V, \|\cdot\|_e) \longrightarrow (V', \|\cdot\|')$ is bounded, $f : (V, \|\cdot\|) \longrightarrow (V', \|\cdot\|')$ is also bounded.

If f is not injective, $\dim(f(V)) < n$, so $(f(V), \|\cdot\|')$ is complete, $f(V)$ is thus closed.

If f is injective, then $\|f(\cdot)\|'$ and $\|\cdot\|$ are equivalent norms on V. So $(f(V), \|\cdot\|')$ is complete. Hence $f(V)$ is closed.

□

Proposition 8.7.20 Let (X, d) be a complete semimetric space. Let $Y \subseteq X$ be a closed subset. Then (Y, d) is complete.

Proof Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in Y. It is also a Cauchy sequence in X, so converges to some $l \in X$. Since Y is closed, one has $l \in Y$. So (Y, d) is complete. □

8.8 Compactness

Definition 8.8.1 Let X be a set and $\mathcal{F}$ be a non-degenerate filter on X. If there does not exists any non-degenerate filter $\mathcal{G}$ on X such that $\mathcal{F} \subsetneq \mathcal{G}$, then we say that $\mathcal{F}$ is an ultrafilter.

Proposition 8.8.2 For any non-degenerate filter $\mathcal{F}$ on X, there exists an ultrafilter on X containing $\mathcal{F}$.

Proof Let Θ be a set of non-degenerate filters containing $\mathcal{F}$, equipped with $\subseteq$. Let Θ_0 be a non-empty totally ordered subset. Let $\mathcal{F}' = \bigcup_{\mathcal{H} \in \Theta_0} \mathcal{H}$.
We prove that $\mathcal{F}'$ is a filter.
(1) Let $(V_1, V_2) \in \mathcal{F}' \times \mathcal{F}'$, $\exists \mathcal{H}_1, \mathcal{H}_2$ in Θ_0, $V_1 \in \mathcal{H}_1,\ V_2 \in \mathcal{H}_2$. Since Θ_0 is totally ordered, either $\mathcal{H}_1 \subseteq \mathcal{H}_2$ or $\mathcal{H}_2 \subseteq \mathcal{H}_1$.
If $\mathcal{H}_1 \subseteq \mathcal{H}_2, V_1 \cap V_2 \in \mathcal{H}_2 \subseteq \mathcal{F}'$. If $\mathcal{H}_2 \subseteq \mathcal{H}_1, V_1 \cap V_2 \in \mathcal{H}_1 \subseteq \mathcal{F}'$.
(2) Let $V \in \mathcal{F}'$, let $\mathcal{H} \in \Theta_0$ such that $V \in \mathcal{H}$. $\forall U \in \mathscr{P}(X), U \supseteq V$, one has $U \in \mathcal{H} \subseteq \mathcal{F}'$. So $\mathcal{F}' \in \Theta$. It is an upper bound of Θ_0. By Zorn's lemma, there exists maximal $\mathcal{G} \in \Theta$, it is an ultrafilter containing $\mathcal{F}$.

□

Proposition 8.8.3 Let X be a set and $\mathcal{F}$ be a non-degenerate filter on X. The following conditions are equivalent.
(1) $\mathcal{F}$ is an ultrafilter.
(2) $\forall A \in \mathscr{P}(X)$, either $A \in \mathcal{F}$ or $X \backslash A \in \mathcal{F}$.
(3) $\forall (A, B) \in \mathscr{P}(X)^2$, if $A \cup B \in \mathcal{F}$, then $A \in \mathcal{F}$ or $B \in \mathcal{F}$.

Proof
(1)⇒(2): Suppose that $A \in \mathscr{P}(X)$ such that $A \notin \mathcal{F}$ and $X \backslash A \notin \mathcal{F}$. Let $B \in \mathcal{F}$. If $B \cap A = \varnothing$, then $B \subseteq X \backslash A$, $(X \backslash A) \in \mathcal{F}$. So $\mathcal{F} \cup \{A\}$ generates a non-degenerate filter $\mathcal{F}' \supsetneq \mathcal{F}$, contradiction.
(2)⇒(3): Suppose that $B \notin \mathcal{F}$. Then $X \backslash B \in \mathcal{F}$. So $(A \cup B) \cap (X \backslash B) \in \mathcal{F}$. So $A \in \mathcal{F}$.
(3)⇒(1): Let $\mathcal{F}'$ be a non-degenerate filter such that $\mathcal{F} \subsetneq \mathcal{F}'$. Take $A \in \mathcal{F}' \backslash \mathcal{F}$. Then $X = A \cup (X \backslash A)$. Since $A \notin \mathcal{F}$, $X \backslash A \in \mathcal{F} \subseteq \mathcal{F}'$. So $\varnothing = A \cap (X \backslash A) \in \mathcal{F}'$. Contradiction. □

Corollary 8.8.4 Let $f : X \longrightarrow Y$ be mapping of sets. If $\mathcal{F}$ is an ultrafilter on X, then $f_*(\mathcal{F})$ is an ultrafilter on Y.

Proof Let A and B be subsets of Y such that

$$A \cup B \in f_*(\mathcal{F}) := \{C \subseteq Y \mid f^{-1}(C) \in \mathcal{F}\}$$

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B) \in \mathcal{F}.$$

Since $\mathcal{F}$ is an ultrafilter, $f^{-1}(A) \in \mathcal{F}$ or $f^{-1}(B) \in \mathcal{F}$. Namely, $A \in f_*(\mathcal{F})$ or $B \in f_*(\mathcal{F})$. □

Definition 8.8.5 Let $(X, \mathscr{T})$ be a topological space, $Y \subseteq X$ and $(U_i)_{i\in I}$ be a family of subset of X.
(1) If $Y \subseteq \bigcup_{i\in I} U_i$, we say that $(U_i)_{i\in I}$ is a **cover** of Y.
(2) If $\exists J \in I$ such that $Y \subseteq \bigcup_{j\in J} U_j$, we say that $(U_j)_{j\in J}$ is a **subcover** of $(U_i)_{i\in I}$ of Y.
(3) If $(U_i)_{i\in I} \in \mathscr{T}^I$ is a cover of Y, we say that ti is an **open cover** of Y.
(4) If I is a finite set and $(U_i)_{i\in I}$ is a cover of Y, we say that $(U_i)_{i\in I}$ is a **finite open cover**.

Proposition 8.8.6 Let $(X, \mathscr{T})$ be a topological space and $Y \subseteq X$. The following conditions are equivalent:
(1) For any ultrafilter $\mathcal{G}$ on X such that $Y \in \mathcal{G}$, $\mathcal{G}$ has a limit point in Y.
(2) For any non-degenerate filter $\mathcal{F}$ on X, such that $Y \in \mathcal{F}$, $\mathcal{F}$ has an adherent point in Y.
(3) Any open cover of Y has a finite subcover.

Proof
(1)⇒(2) Let $\mathcal{G}$ be an ultrafilter containing $\mathcal{F}$ and $x \in Y$ be a limit point of $\mathcal{G}$. For any $U \in \mathcal{V}_x(\mathscr{T})$, $U \in \mathcal{G}$, so $\forall A \in \mathcal{F}$, $U \cap A \neq \varnothing$.
(2)⇒(1) Let $\mathcal{G}$ be an ultrafilter, and let $x \in Y$ be an adherent point of $\mathcal{G}$. Then the filter $\mathcal{G}'$ generated by $\mathcal{G} \cup \mathcal{V}_x(\mathscr{T})$ is non-degenerate and contains $\mathcal{G}$. So $\mathcal{G}' = \mathcal{G}$. This means $\mathcal{V}_x(\mathscr{T}) \subseteq \mathcal{G}$.
(2)⇒(3) Let $(U_i)_{i\in I}$ be an open cover of Y. Suppose that it does not have any finite subcover. For any $i \in I$, let $F_i = X \backslash U_i$. For any finite subset J of I,

$Y \nsubseteq \bigcup_{j\in J} U_j$. So

$$Y \bigcap \left(X \setminus \bigcup_{j\in J} U_j \right) = Y \bigcap \left(\bigcap_{j\in J} F_j \right) \neq \varnothing.$$

So $\{F_i \mid i \in I\} \cup \{Y\}$ generates a non-degenerate filter $\mathcal{F}$. It has an adherent point of $x \in Y$. Since $Y \subseteq \cup_{i\in I} U_i$, $\exists i_0 \in I$, $x \in U_{i_0}$. So $U_{i_0} \in \mathcal{V}_x(\mathscr{T})$. This is impossible since $U_{i_0} \cap F_{i_0} = \varnothing$.
(3)⇒(2) Let $\mathcal{F}$ be a non-degenerate filter such that $Y \in \mathcal{F}$. Suppose that $\mathcal{F}$ does not have any adherent point in Y.

For any $y \in Y$, there exists open neighborhood U_y of Y and $A_y \in \mathcal{F}$, such that $U_y \cap A_y = \varnothing$. Since $Y \subseteq \bigcup_{y\in Y} U_y$, $\exists\{y_1, \ldots, y_n\} \subseteq Y$ such that $Y \subseteq \bigcup_{i=1}^{n} U_{y_i}$. Take $A = \left(\bigcap_{i=1}^{n} A_{y_i} \right) \cap Y \in \mathcal{F}$, $A \neq \varnothing$, $A \subseteq Y$.

$$\begin{aligned} A = A \cap Y &\subseteq A \bigcap \bigcup_{i=1}^{n} U_{y_i} \\ &= \bigcup_{i=1}^{n} (A \cap U_{y_i}) \\ &\subseteq \bigcup_{i=1}^{n} (A_{y_i} \cap U_{y_i}) \\ &= \varnothing. \end{aligned}$$

Contradiction. □

Definition 8.8.7 Let $(X, \mathscr{T})$ be a topological space. If $Y \subseteq X$ satisfies the equivalent conditions described in the previous proposition, we say that Y is **compact**.

Proposition 8.8.8 Let $(X, \mathscr{T}_X)$ and $(Y, \mathscr{T}_Y)$ be two topological spaces, $f : X \longrightarrow Y$ be a continuous mapping. If $F \subseteq X$ is compact, then $f(F)$ is also compact.

Proof Let $(V_i)_{i\in I}$ be a open cover of $f(F)$. Then

$$F \subseteq f^{-1}(f(F)) \subseteq \bigcup_{i\in I} f^{-1}(V_i).$$

Since F is compact, $\exists J \subseteq I$ finite such that $F \subseteq \bigcup_{j \in J} f^{-1}(V_j)$. Hence $f(F) \subseteq \bigcup_{j \in J} V_j$. □

Proposition 8.8.9 Let $(X, \mathscr{T})$ be a topological space, A be a compact subset of X, F is a closed subset of X, then $A \cap F$ is a compact subset of X.

Proof Let $(U_i)_{i \in I}$ be an open cover of $A \cap F$. Then

$$A \subseteq \left(\bigcup_{i \in I} U_i \right) \cup (X \backslash F) .$$

So $\exists J \subseteq I$ finite, $A \subseteq (\bigcup_{j \in J} U_j) \cup (X \backslash F)$. Hence

$$A \cap F \subseteq \left(\bigcup_{j \in J} U_j \right) \cup ((X \backslash F) \cap F) = \bigcup_{j \in J} U_j .$$

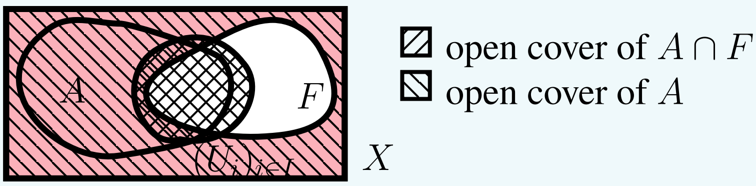


Figure 8.8.1: proposition 8.8.9

□

Proposition 8.8.10 Let $(X, \mathscr{T})$ be a Hausdorff topological space, A be a compact subset of X.
(1) For any $x \in X \backslash A$, there exists open subsets U and V of X, such that $A \subseteq U$, $x \in V$ and $U \cap V = \varnothing$.
(2) A is closed.

Proof
(1) $\forall y \in A$, $\exists U_y \in \mathscr{T}$, $V_y \in \mathscr{T}$ such that $y \in U_y$, $x \in V_y$, $U_y \cap V_y = \varnothing$. (Hausdorff) Since $A \subseteq \bigcup_{y \in A} U_y$ and A is compact. $\exists \{y_1, \ldots, y_n\} \subseteq A$, $A \subseteq$

$\bigcup_{i=1}^{n} U_{y_i}$. Let $U = \bigcup_{i=1}^{n} U_{y_i}$, $V = \bigcap_{i=1}^{n} V_{y_i}$. These are open subsets of X, and $A \subseteq U$, $x \in V$.

$$U \cap V = \bigcup_{i=1}^{n} U_{y_i} \cap V \subseteq \bigcup_{i=1}^{n} U_{y_i} \cap V_{y_i} = \varnothing.$$

(1)⇒(2): $\forall x \in X \backslash A$, $\exists (U, V) \in \mathscr{T}^2$, $A \subseteq U$, $x \in V$, $U \cap V = \varnothing$. Hence, $V \subseteq X \backslash A$. So $X \backslash A$ is a neighborhood of x.

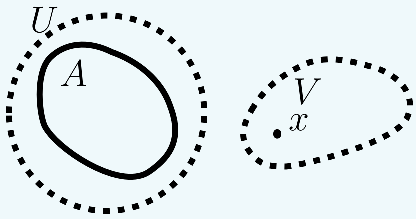


Figure 8.8.2: proposition 8.8.10

□

Corollary 8.8.11 Let $(X, \mathscr{T})$ be a Hausdorff topological space, and A and B be disjoint compact subset of X. There exist $(U, V) \in \mathscr{T}^2$, such that $A \subseteq U$, $B \subseteq V$ and $U \cap V = \varnothing$.

Proof By (1) of the previous proposition.

$$\forall x \in B,\ \exists (U_x, V_x) \in \mathscr{T}^2,\ A \subseteq U_x,\ x \in V_x,\ U_x \cap V_x = \varnothing.$$

$B \subseteq \bigcup_{x \in B} V_x$. So, $\exists \{x_1, \ldots, x_n\} \subseteq B$, such that $B \subseteq \bigcup_{i=1}^{n} V_{x_i}$. Let

$$U = \bigcap_{i=1}^{n} U_{x_i} \in \mathscr{T},\ V = \bigcup_{i=1}^{n} V_{x_i} \in \mathscr{T}.$$

Then, $A \subseteq U$, $B \subseteq V$, and $U \cap V = \varnothing$. □

Theorem 8.8.12 Let $(X, \mathscr{T})$ be a topological space and A be a compact subset of X. Let $(I, \leq)$ be a partially ordered set and $(F_i)_{i \in I}$ be a decreasing family of closed subsets of X. Assume that, any finite subset of I has an upper bound in I. If $\left(\bigcap_{i \in I} F_i\right) \cap A = \varnothing$, then exists $i_0 \in I$ such that $F_{i_0} \cap A = \varnothing$.

Remark 8.8.13 In the particular case where $I = \mathbb{N}$, the theorem becomes: If $(A_n)_{n\in\mathbb{N}}$ is a sequence of compact non-empty subsets of X such that $A_0 \supseteq A_1 \supseteq A_2 \supseteq \cdots$, then $\bigcap\limits_{n\in\mathbb{N}} A_n \neq \varnothing$.

Proof $\forall i \in I$, let $U_i = X \backslash F_i \in \mathscr{T}$. Since $\left(\bigcap\limits_{i\in I} F_i\right) \cap A = \varnothing$. $A \subseteq X \backslash \bigcap\limits_{i\in I} F_i = \bigcup\limits_{i\in I} U_i$. Since A is compact, $\exists J \subseteq I$, finite such that $A \subseteq \bigcup\limits_{j\in J} U_j$. Let i_0 be an upper bound of J. $\forall j \in J$, $j \leq i_0$. So $F_j \supset F_{i_0}$. $U_j \subseteq U_{i_0}$. Hence $A \subseteq U_{i_0}$, $A \cap F_{i_0} = \varnothing$. □

Theorem 8.8.14 (Tychonoff) Let I be a non-empty set and $(X_i, \mathscr{T}_i)$, $i \in I$ be topological spaces. Let $X = \prod\limits_{i\in I} X_i$, and $\mathscr{T}$ be the product topology of $(\mathscr{T}_i)_{i\in I}$. For any $i \in I$, let A_i be a compact subset of X_i, let $A = \prod\limits_{i\in I} A_i \subseteq X$. Then A is compact subset of X.

Proof For any $i \in I$, let $\pi_i : (x_j)_{j\in I} \longmapsto x_i$ be the projection mapping.

$$A = \bigcap_{i\in I} \pi_i^{-1}(A_i).$$

Let $\mathcal{F}$ be an ultrafilter on X such that $A \in \mathcal{F}$. Then $\forall i \in I$, $\pi_{i*}(\mathcal{F}) =: \mathcal{F}_i$ is an ultrafilter on X_i such that $A_i \in \mathcal{F}_i$. Since A_i is compact, $\mathcal{F}_i$ has a limit point $x_i \in A_i$. Let $x = (x_i)_{i\in I}$. Let U be a neighborhood of x. There exists $J \subseteq I$ finite and $U_j \in J_j$ $(j \in J)$ such that $x_j \in U_j$ and $\bigcap\limits_{j\in J} \pi_j^{-1}(U_j) \subseteq U$. $\forall j \in J, U_j \in \mathcal{F}_j$. Since x_j is a limit point of $\mathcal{F}_j$. Hence $\pi_j^{-1}(U_j) \in \mathcal{F}$. Thus $\bigcap\limits_{j\in J} \pi_j^{-1}(U_j) \in \mathcal{F}$, which implies $U \in \mathcal{F}$. □

8.9 Compact Metric Spaces

Definition 8.9.1 Let $(X, \mathscr{T})$ be a topological space and $Y \subseteq X$. If any sequence in Y has a subsequence that converges to some element of X. We say that Y is **sequentially compact**.

Theorem 8.9.2 Let $(X, \mathscr{T})$ be a topological space and $Y \subseteq X$. If Y is compact and if each $y \in Y$ has a countable neighborhood basis, then Y is sequentially compact.

Proof Let $x = (x_n)_{n\in I}$ be a sequence in Y and $\mathcal{F}$ be the Fréchet filter of I. Then $x_*(\mathcal{F})$ is a non-degenerate filter on X and $Y \in x_*(\mathcal{F})$. Since Y is compact, $x_*(\mathcal{F})$ has an adherent point $p \in Y$. Let $(U_k)_{k\in\mathbb{N}}$ be a decreasing sequence of neighborhood of p such that $\{U_k \mid k \in \mathbb{N}\}$ forms a neighborhood basis of p. Therefore we can construct in a recursive way a strictly increasing sequence $(n_k)_{k\in\mathbb{N}}$ in I such that $x_{n_k} \in U_k$. Thus $(x_{n_k})_{k\in\mathbb{N}}$ converges to p. $\square$

Theorem 8.9.3 Let (X, d) be a metric space, and $Y \subseteq X$. The following conditions are equivalent:
(1) Y is compact.
(2) (Y, d) is complete and

$$\forall \varepsilon > 0,\ \exists A_\varepsilon \subseteq Y \text{ finite},\ Y \subseteq \bigcup_{x\in A_\varepsilon} B(x, \varepsilon).$$

(3) Y is sequentially compact.

Proof
(1)$\Rightarrow$ (2) Let $\mathcal{F}$ be a Cauchy filter on Y. Let $f : Y \longrightarrow X$ be the inclusion mapping. Then $f_*(\mathcal{F})$ is a Cauchy filter on X and $Y \in f_*(\mathcal{F})$. Since Y is compact, $f_*(\mathcal{F})$ has an adherent point $l \in Y$. So l is a limit point of $f_*(\mathcal{F})$ (since $f_*(\mathcal{F})$ is a Cauchy filter.) For any $U \in \mathcal{V}_l(\mathscr{T})$, $U \in f_*(\mathcal{F})$, namely, $f^{-1}(U) = U \cap Y \in \mathcal{F}$. Thus l is a limit point of $\mathcal{F}$. Since $Y \subseteq \bigcup_{y\in Y} B(y, \varepsilon)$. Since Y is compact, $\exists A_\varepsilon \subseteq Y$ finite, such that $Y \subseteq \bigcup\limits_{x\in A_\varepsilon} B(x, \varepsilon)$.
(2)$\Rightarrow$ (1) Let $\mathcal{F}$ be an ultrafilter on X such that $Y \in \mathcal{F}$. For any $\varepsilon > 0$, $\bigcup\limits_{x\in A_\varepsilon} B(x, \varepsilon) \in \mathcal{F}$. Hence $\mathcal{F}$ is a Cauchy filter, which implies that $\mathcal{F}|_Y := \{Y \cap U \mid U \in \mathcal{F}\}$ is a Cauchy filter. Thus $\mathcal{F}|_Y$ has a limit point $l \in Y$, which is also a limit

point of $\mathcal{F}$.
(2)$\Rightarrow$ (3) is already known.
(3)$\Rightarrow$ (1) Let $(x_n)_{n\in I}$ be a Cauchy sequence in Y. Since Y is sequentially compact, $(x_n)_{n\in I}$ has a subsequence that converges to some $l \in Y$. Since $(x_n)_{n\in I}$ is a Cauchy sequence, it converges to l. We prove the second statement by contradiction.

Assume that $\varepsilon > 0$ is such that Y cannot be covered by finitely many balls centered in Y with radius ε. We construct recursively a sequence $(x_n)_{n\in\mathbb{N}}$ as follows. $x_0 \in Y$ is chosen arbitrarily. If $x_0, \ldots, x_n$ are chosen, we pick

$$x_{n+1} \in Y \backslash \bigcup_{i=0}^{n} B(x_i, \varepsilon).$$

Then $\forall (i,j) \in \mathbb{N}^2, i \neq j, \mathrm{d}(x_i, x_j) < \varepsilon$. So $(x_n)_{n\in\mathbb{N}}$ does not have any Cauchy subsequence. It cannot have a convergent subsequence. □

Example 8.9.4 Let $(a,b) \in \mathbb{R}^2$, $a < b$. $[a,b]$ is closed. $\forall \varepsilon, \exists N \in \mathbb{N}_{>0}. \frac{b-a}{N} < \varepsilon$. $\forall i \in \{0, \ldots, N\}$, let $x_i = a + \frac{i}{N}(b-a)$. $a = x_0 < x_1 < \cdots < x_{N-1} < x_N = b$.

$$[a,b] \subseteq \bigcup_{i=0}^{N} B(x_i, \varepsilon) = \bigcup_{i=0}^{n}]x_i - \varepsilon, x_i + \varepsilon[.$$

So $[a,b]$ is compact.

Proposition 8.9.5 Let (X, d) be a metric space and $Y \subseteq X$. If Y is compact, then Y is bounded and closed.

Proof Since X is Hausdorff, Y is closed. $\exists A \subseteq Y$ finite, $Y \subseteq \cup_{x\in A} B(x, 1)$. Since each $B(x,1)$ is bounded, so is Y. □

Definition 8.9.6 Let $(X, \mathscr{T})$ be a topological space. If $\forall x \in X$, x has a compact neighborhood, we say that X is **locally compact**.

Example 8.9.7 $(\mathbb{R}), |\cdot|$ is locally compact.

$$\forall x \in \mathbb{R}, [x-1, x+1] \text{ is compact neighborhood of } x.$$

Proposition 8.9.8 Let $(K, |\cdot|)$ be a locally compact valued field. Then,
(1) $(K, |\cdot|)$ is complete.
(2) Assume that exists $a \in K$, $|a| > 1$. Let $(V, \|\cdot\|)$ be a finite-dimensional vector space over K. Then any bounded closed subset of V is compact.
(3) $(V, \|\cdot\|)$ is locally compact.

Proof
(2) Assume that $V = K^n$, $\|(a_1, \ldots, a_n)\| = \max\{|a_1|, \ldots, |a_n|\}$. Let $A \subseteq K^n$, bounded, closed. Let $R > 0$, $A \subseteq \bar{B}(0_V, R)$.

Since $(K, |\cdot|)$ is locally compact, there exists a compact neighborhood U of $0_K \in K$. $\exists \varepsilon > 0, \bar{B}(0, \varepsilon) \subseteq U$. So $\bar{B}(0_K, \varepsilon)$ is compact. Take $n \in \mathbb{N}_{\geq 1}$ such that $|a|^n \varepsilon > R$.

$$f : K \longrightarrow K \text{continuous}$$

$$b \longmapsto a^n b$$

$f\left(\bar{B}(0_K, \varepsilon)\right) = a^n \bar{B}(0_K, \varepsilon) \supseteq \bar{B}(0_K, R)$. is compact. By Tychonoff theorem, $\bar{B}(0_V, R)$ is compact. Since A is closed, A is compact.
(3) $\bar{B}(0_V, R)$ is compact for any $R > 0$.
(1) $B(0_K, R)$ is compact for any $R > 0$. For any Cauchy sequence $(a_n)_{n\in\mathbb{N}}$ in K, $\exists R > 0$. $\{a_n \mid n \in \mathbb{N}\} \subseteq \bar{B}(0_K, R)$, so $(a_n)_{n\in\mathbb{N}}$ converges. □

Lemma 8.9.9 Let $A \subseteq R$ be a non-empty compact subset. Then A has a greatest and least element.

Proof Since A is non-empty and bounded, $\{\sup A, \inf A\} \subseteq \mathbb{R}$. Since $A = \bar{A}$, $\sup A \in A, \inf A \in A$. □

Theorem 8.9.10 Let $(X, \mathscr{T})$ be a topological space and $f : X \longrightarrow \mathbb{R}$ be a continuous mapping. If $Y \subseteq X$ is compact, then $f|_Y$ has its maximum and minimum. ($\exists x_1 \in Y, f(x_1) \geq f(y), \forall y \in Y$. $\exists x_2 \in Y, \forall y \in Y, f(x_2) \leq f(y)$.)

Proof $f(Y) \subseteq \mathbb{R}$ is compact. □

Theorem 8.9.11 Let (X, d_X) and (Y, d_Y) be metric spaces, $f : X \longrightarrow Y$ be a continuous mapping. If (X, d_X) is compact, then f is uniformly continuous.

Proof f is uniformly continuous if and only if

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x, y \in X,\ \mathrm{d}_X(x, y) < \delta \Rightarrow \mathrm{d}_Y(f(x), f(y)) < \varepsilon.$$

Suppose by contradiction that f is not uniformly continuous. That is

$$\exists \varepsilon > 0,\ \forall \delta > 0,\ \exists (x_1, x_2) \in X^2,\ \mathrm{d}_X(x_1, x_2) < \delta \text{ and } \mathrm{d}_Y(f(x_1), f(x_2)) \geq \varepsilon.$$

So we can choose sequences $(x_n)_{n \in \mathbb{N}_{\geq 1}}$ and $(y_n)_{n \in \mathbb{N}_{\geq 1}}$ in X such that

$$\forall n \in \mathbb{N}, \mathrm{d}_X(x_n, y_n) \leq \frac{1}{n},\ \mathrm{d}_Y(f(x_n), f(y_n)) \geq \varepsilon.$$

By compactness of X, there exists a subsequence $I \subseteq \mathbb{N}_{\geq 1}$ finite, such that $(x_n)_{n \in I}$ converges to some $a \in X$, and $(y_n)_{n \in I}$ converges to some $b \in X$. Since $\mathrm{d}_X : X \times X \longrightarrow \mathbb{R}_{\geq 0}$ is continuous,

$$0 \leq \mathrm{d}_X(a, b) = \lim_{n \in I, n \to +\infty} \mathrm{d}_X(x_n, y_n) \leq \lim_{n \in I, n \to +\infty} \frac{1}{n}.$$

$\square$

8.10 Path Connectedness

Definition 8.10.1 Let $(X, \mathscr{T})$ be a topological space and C be a subset of X. If for any $(x, y) \in C \times C$, there exists a continuous mapping $\varphi : [0, 1] \longrightarrow C$, such that $\varphi(0) = x$ and $\varphi(1) = y$, we say that C is **path connected**.

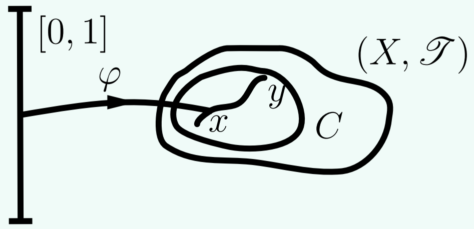


Figure 8.10.1: Path connected

Proposition 8.10.2 Let $(X, \mathscr{T}_X)$, $(Y, \mathscr{T}_Y)$ be two topological spaces and $f : X \longrightarrow Y$ be continuous mapping. If $C \subseteq X$ is path connected, then $f(C) \subseteq Y$ is path connected.

Proof Let $(a, b) \in C \times C$. There exist $(x, y) \in C \times C$ such that $f(x) = a$ and $f(y) = b$. Since C is path connected, so we have a continuous $\varphi : [0, 1] \longrightarrow C$ such that $\varphi(0) = x$ and $\varphi(1) = y$. Then $f \circ \varphi : [0, 1] \longrightarrow f(C)$ is continuous, and

$$(f \circ \varphi)(0) = f(x) = a, \ (f \circ \varphi)(1) = f(y) = b.$$

Hence $f(C)$ is path connected. □

Theorem 8.10.3 Let I be a subset of $\mathbb{R}$. I is path connected if and only if I is an interval.

Proof
$\Leftarrow$ Assume that I is an interval. Let $(a, b) \in I \times I$ such that $a \leq b$. One has $[a, b] \subseteq I$. Consider $\varphi : [0, 1] \longrightarrow I$ defined as:

$$\varphi(t) = \frac{a+b}{2} + \frac{b-a}{2}t, \ t \in [0, 1].$$

This is the sum of a constant mapping and a linear mapping. Hence φ is continuous. So I is path connected.
$\Rightarrow$ Assume that I is path connected. Let $(a, b) \in I \times I$, $a \leq b$, $\varphi : [0, 1] \longrightarrow I$ a continuous mapping such that $\varphi(0) = a$ and $\varphi(1) = b$. Let $c \in [a, b]$. It is

suffices to show that $c \in I$. Assume that $c \notin I$. We have $a < c < b$. We will define two sequences $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ in a recursive way: $x_0 = 0, y_0 = 1$. We want to have $\varphi(x_n) < \varphi(y_n)$. Let $(z_n) := \frac{x_n+y_n}{2}$.

$$\begin{cases} x_{n+1} = x_n, y_{n+1} = z_n, \text{ if } \varphi(z_n) > c, \\ x_{n+1} = z_n, y_{n+1} = y_n, \text{ if } \varphi(z_n) < c. \end{cases}$$

Then the sequence $(x_n)_{n\in\mathbb{N}}$ is increasing, $(y_n)_{n\in\mathbb{N}}$ is decreasing, and $0 \leq y_n - x_n \leq \frac{1}{2^n}$. So $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ converge to the same point $d \in [0,1]$. Note that

$$\forall n \in \mathbb{N},\ (c - \varphi(x_n))\,(c - \varphi(y_n)) < 0.$$

$$0 \leq (c - \varphi(d))^2 = \lim_{n\to\infty} (c - \varphi(x_n))\,(c - \varphi(y_n)) \leq 0.$$

So, $\varphi(d) = c$. Contradiction. □

Corollary 8.10.4 Let I be an interval in $\mathbb{R}$. If $f : I \longrightarrow R$ is a continuous mapping, then $f(I)$ is also an interval. In particular, if f has a positive value and a negative value, then it must have a zero.

Proof Since I is path connected, so $f(I)$ is path connected. Hence $f(I) \subseteq \mathbb{R}$ is an interval. □

Chapter 9

Differential Calculus

9.1 Landau symbol

In this section, we fix a complete valued field $(K, |\cdot|)$ and a normed vector space $(V, \|\cdot\|)$ over K.

Definition 9.1.1 Let X be a set, $f : X \longrightarrow V$, $g : X \longrightarrow \mathbb{R}_{\geq 0}$ be mappings. Let $Y \subseteq X$ be a subset. We use the expression

$$f(x) = \mathcal{O}(g(x))$$

to denote the statement:

$$\exists C > 0, \forall x \in Y, \|f(x)\| \leq C \cdot g(x).$$

Let $\mathcal{F}$ be a filter on X, we use the expression

$$f(x) = \mathcal{O}\left(g(x)\right) \text{ along } \mathcal{F}$$

to denote the statement:

$$\exists C > 0, \exists A \in \mathcal{F}, \|f(x)\| \leq C \cdot g(x), \forall x \in A.$$

We use the expression

$$f(x) = o\left(g(x)\right) \text{ along } \mathcal{F}$$

to denote the statement:

$$\exists \varepsilon : X \longrightarrow \mathbb{R}_{\geq 0}, \exists A \in \mathcal{F}, \lim_{\mathcal{F}} \varepsilon = 0 \text{ and } \forall x \in A, \|f(x)\| \leq \varepsilon(x) g(x).$$

Proposition 9.1.2 Let X be a set and $\mathcal{F}$ be a filter on X.
(1) Let $f : X \longrightarrow V, g : X \longrightarrow \mathbb{R}_{\geq 0}$ be mappings. If $f(x) = o(g(x))$ along $\mathcal{F}$, then $f(x) = \mathcal{O}(g(x))$ along $\mathcal{F}$.
(2)

1. Let $f_1 : X \longrightarrow V$, $f_2 : X \longrightarrow V$ and $g : X \longrightarrow \mathbb{R}_{\geq 0}$ be mappings. If $f_1(x) = \mathcal{O}(g(x))$ and $f_2(x) = \mathcal{O}(g(x))$ along $\mathcal{F}$, then $f_1(x) + f_2(x) = \mathcal{O}(g(x))$ along $\mathcal{F}$.

2. Let $f_1 : X \longrightarrow V$, $f_2 : X \longrightarrow V$ and $g : X \longrightarrow \mathbb{R}_{\geq 0}$ be mappings. If $f_1(x) = o(g(x))$ and $f_2(x) = o(g(x))$ along $\mathcal{F}$, then $f_1(x)+f_2(x) = o(g(x))$ along $\mathcal{F}$.

(3) Let $\lambda : X \longrightarrow K$, $f : X \longrightarrow V$, $g : X \longrightarrow \mathbb{R}_{\geq 0}$, $h : X \longrightarrow \mathbb{R}_{\geq 0}$ be mappings.

1. If $\lambda(x) = \mathcal{O}\left(g(x)\right)$ along $\mathcal{F}$, $f(x) = \mathcal{O}\left(h(x)\right)$ along $\mathcal{F}$, then

$$(\lambda f)(x) = \lambda(x) f(x) = \mathcal{O}\left(g(x)h(x)\right).$$

2. If $\lambda(x) = \mathcal{O}\left(g(x)\right)$ along $\mathcal{F}$, $f(x) = o\left(h(x)\right)$ along $\mathcal{F}$, or if $\lambda(x) = o\left(g(x)\right)$ along $\mathcal{F}$, $f(x) = \mathcal{O}\left(h(x)\right)$ along $\mathcal{F}$, then

$$\lambda(x) f(x) = o\left(g(x)h(x)\right).$$

Proof
(1) We have $\varepsilon : X \longrightarrow \mathbb{R}_{\geq 0}$, $A \in \mathcal{F}$ such that $\lim_{\mathcal{F}} \varepsilon = 0$ and $\forall x \in A$, $\|f(x)\| \leq \varepsilon(x) g(x)$. Since $\lim_{\mathcal{F}} \varepsilon = 0$, there exists $B \in \mathscr{T}$ such that $\forall x \in B$, $|\varepsilon(x)| < 1$, hence $\forall x \in A \cap B$, $\|f(x)\| \geq (x)$.
(2)

1. $A_1, A_2 \in \mathcal{F}, C_1, C_2 > 0$, $\forall x \in A_1, \|f_1(x)\| \leq C_1 g(x), \forall x \in A_2, \|f_2(x)\| \leq C_2 g(x)$. So $f_1(x) + f_2(x) = \mathcal{O}(g(x))$

2. Let $\varepsilon_1 : X \longrightarrow \mathbb{R}_{\geq 0}$, $\varepsilon_2 : X \longrightarrow \mathbb{R}_{\geq 0}$, $A \in \mathcal{F}$, $\lim_{\mathcal{F}} \varepsilon_1 = \lim_{\mathcal{F}} \varepsilon_2 = 0$. $\forall x \in A_1, \|f_1(x)\| \leq \varepsilon_1(x) \cdot g(x), \forall x \in A_2, \|f_2(x)\| \leq \varepsilon_2(x) g(x)$. So $\lim_{\mathcal{F}} \varepsilon_1 + \varepsilon_2 = 0$.

$$\forall x \in A_1 \cap A_2,\ \|f_1(x) + f_2(x)\| \leq \|f_1(x)\| + \|f_2(x)\| \leq \left(\varepsilon_1(x) + \varepsilon_2(x)\right) g(x).$$

(3)

1. There exists $(C_1, C_2) \in \mathbb{R}^2_{>0}$ and $(A_1, A_2) \in \mathcal{F}^2$ such that

$$\forall x \in A_1,\ |\lambda(x)| \leq C_1 g(x),\ \forall x \in A_2,\ \|f(x)\| \leq C_2 h(x).$$

Hence,

$$\forall x \in A_1 \cap A_2,\ \|(\lambda(x) f(x))\| \leq |\lambda(x)| \cdot \|f(x)\| \leq C_1 C_2 g(x) h(x).$$

2. We assume that

$$\lambda(x) = \mathcal{O}\left(g(x)\right) \text{ along } \mathcal{F},\ f(x) = o\left(h(x)\right) \text{ along } \mathcal{F}.$$

There exists $(A_1, A_2) \in \mathcal{F} \times \mathcal{F}, C \in \mathbb{R}_{\geq 0}$ and a mapping $\varepsilon : X \longrightarrow \mathbb{R}_{\geq 0}$ such that

$$\forall x \in A_1,\ |\lambda(x)| \leq C \cdot g(x),\ \forall x \in A_2, \|f(x)\| \leq \varepsilon(x) h(x).$$

Then one has

$$\lim_{\mathcal{F}} C\varepsilon(x) = 0$$

and

$$\forall x \in A_1 \cap A_2,\ \|(\lambda(x) f(x))\| \leq |\lambda(x)| \cdot \|f(x)\| \leq C \cdot g(x) \cdot \varepsilon(x) h(x)$$

As required.

□

Example 9.1.3
(1) Let $I \subseteq \mathbb{N}$ infinite. Let $(V, \|\cdot\|)$ be a normed vector space over complete valued field $(K, |\cdot|)$. Let $\mathcal{F}$ be the filter on I. Let $(x_n)_{n\in I} \in V^I, (b_n)_{n\in I} \in \mathbb{R}^I_{\geq 0}$. We denote by

$$x_n = \mathcal{O}\left(b_n\right),\ n \in I,\ n \to +\infty$$

the statement $x_n = \mathcal{O}\left(b_n\right)$ along $\mathcal{F}$. Namely,

$$\exists N \in \mathbb{N},\ \exists C > 0,\ \forall n \in I_{\geq N},\ \|x_n\| \leq C \cdot b_n.$$

$$x_n = o\left(b_n\right),\ n \in I,\ n \to +\infty$$

denotes the statement $x_n = o\left(b_n\right)$ along $\mathcal{F}$. Namely,

$$\exists (\varepsilon_n)_{n\in I} \text{ such that } \lim_{n\to+\infty} \varepsilon_n = 0,\ \exists N \in \mathbb{N},\ \forall n \in I_{\geq N},\ \|x_n\| \leq \varepsilon_n \cdot b_n.$$

(2) Let $(X, \mathscr{T})$ be a topological space, $Y \subseteq X$, $y_0 \in \overline{Y}$. Let $f : Y \longrightarrow V$ and $g : Y \longrightarrow \mathbb{R}_{\geq 0}$ be mappings.

$$\mathcal{F} = \mathcal{V}_{y_0}(\mathscr{T})|_Y := \{U \cap Y \mid U \text{ is a neighborhood of } y_0\}$$

$f(y)\mathcal{O}(g(y))$, $y \in Y$, $y \to y_0$ denotes $f(y) = \mathcal{O}(g(y))$ along $\mathcal{F}$. Namely,

$$\exists C > 0,\ \exists U \in \mathcal{V}_{y_0}(\mathscr{T}),\ \forall y \in U \cap Y,\ \|f(y)\| \leq C \cdot g(y).$$

$$f(y) = o(g(y)),\ y \in Y,\ y \to y_0$$

denotes $f(y) = o(g(y))$ along $\mathcal{F}$. Namely,

$$\exists \varepsilon : Y \longrightarrow \mathbb{R}_{\geq 0},\ \lim_{y \in Y, y \to y_0} \varepsilon(y) = 0,\ \exists U \in \mathcal{V}_{y_0}(\mathscr{T}),$$

$$\forall y \in U \cap Y,\ \|f(y)\| \leq \varepsilon(y) g(y).$$

(3) Let $\mathcal{F}$ be a filter on $\mathbb{R}$ generated by subsets of the form $[a, +\infty[$. Let $Y \subseteq \mathbb{R}$ not bounded from above. Let $f : Y \longrightarrow V$ and $g : Y \longrightarrow \mathbb{R}_{\geq 0}$ be mappings. Then

$$f(y) = \mathcal{O}(g(y)),\ y \in Y,\ y \to +\infty$$

denotes $f(y) = \mathcal{O}(g(y))$ along $\mathcal{F}|_Y$. Namely,

$$\exists C > 0,\ \exists a \in \mathbb{R},\ \forall y \in Y_{\geq a},\ \|f(y)\| \leq C \cdot g(y).$$

$$f(y) = o(g(y)),\ y \in Y,\ y \to +\infty$$

denotes $f(y) = o(g(y))$ along $\mathcal{F}|_Y$. Namely,

$$\exists \varepsilon : Y \longrightarrow \mathbb{R}_{\geq 0},\ \lim_{y \to +\infty} \varepsilon(y) = 0,\ \exists a \in \mathbb{R},\ \forall y \in Y_{\geq a},\ \|f(y)\| \leq \varepsilon(y) g(y).$$

9.2 Differentiability

We fix a complete valued field $(K, |\cdot|)$. We suppose that there exists $a \in K^\times$, such that $|a| < 1$. Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over K.

$$\mathscr{L}(E, F) := \{\varphi \in \mathrm{Hom}_K(E, F) \mid \|\varphi\| < +\infty\}.$$

$(\mathscr{L}(E, F), \|\cdot\|)$ is a normed vector space over K.

Definition 9.2.1 Let $U \subseteq E$ be subset and $p \in U^\circ$. We say that a mapping $f : U \longrightarrow F$ is **differentiable** at p if there exists $\varphi \in \mathscr{L}(E, F)$ such that

$$f(p+h) - f(p) - \varphi(h) = o\left(\|h\|_E\right),\ h \to 0_E.$$

If $U = U^\circ$ and f is differentiable at every point of U, we say that f is **differentiable** on U.

Proposition 9.2.2 Assume that $f : U \longrightarrow F$ is differentiable at $p \in U^\circ$. There exists a unique $\varphi \in \mathscr{L}(E, F)$ such that

$$f(p+h) - f(p) - \varphi(h) = o\left(\|h\|_E\right),\ h \to 0_E.$$

Lemma 9.2.3 $\forall \eta \in \mathscr{L}(E, F), \forall r > 0.$

$$\|\eta\| = \sup_{x \in E, 0 < \|x\|_E \leq r} \frac{\|\eta(x)\|_F}{\|x\|_E} = \sup_{\substack{x \in E \\ 0 < \|x\|_E < r}} \frac{\|\eta(x)\|_F}{\|x\|_E}.$$

Proof (of Lemma) $\|\eta\| \geq \sup_{x \in E, 0 < \|x\|_E < r} \frac{\|\eta(x)\|_F}{\|x\|_E}$. $\forall y \in E \backslash \{0\}$, $\|a^N y\|_E = |a|^N \|y\|_E < r$.

$$\frac{\|\eta(a^N y)\|_F}{\|a^N y\|_E} = \frac{|a|^N \cdot \|\eta(y)\|_F}{|a|^N \cdot \|y\|_E} = \frac{\|\eta(y)\|_F}{\|y\|_E} \leq \sup_{\substack{x \in E \\ 0 < \|x\|_E < r}} \frac{\|\eta(x)\|_F}{\|x\|_E}.$$

□

Proof (of Proposition) Suppose $\varphi, \psi \in \mathscr{L}(E, F)$ are such that

$$f(p+h) - f(p) - \varphi(h) = o\left(\|h\|_E\right),\ h \to 0_E,$$

$$f(p+h) - f(p) - \psi(h) = o\left(\|h\|_E\right),\ h \to 0_E.$$

Then

$$\varphi(h) - \psi(h) = o\left(\|h\|_E\right),\ h \to 0_E.$$

$$\exists r > 0, \exists \varepsilon : \overline{B}(0_E, r) \longrightarrow \mathbb{R}_{\geq 0} \text{ such that } \lim_{h \to 0_E} \varepsilon(h) = 0.$$

$$\forall h \in \overline{B}(0_E, r),\ \|(\varphi - \psi)(h)\|_F = \varepsilon(h)\|h\|_E.$$

$$\|\varphi - \psi\| = \sup_{\substack{x \in E \\ 0 < \|h\|_E < r'}} \frac{\|\varphi(h) - \psi(h)\|_F}{\|h\|_E} \leq \sup_{0 < \|h\|_E < r'} \varepsilon(h).$$

Taking the limit when $r' \to 0$, by $\limsup_{h \to 0_E} \varepsilon(h) = 0$. We get $\|\varphi - \psi\| = 0$, hence $\varphi = \psi$. $\square$

Definition 9.2.4 Let $U \subseteq E$ and $f : U \longrightarrow F$ be a mapping that is differentiable at $p \in U^\circ$. The unique $\varphi \in \mathscr{L}(E, F)$ such that

$$f(p+h) - f(p) - \varphi(h) = o\left(\|h\|_E\right), \ h \to 0_E$$

is called the **differential** of f at p and is denoted as

$$\mathrm{D}(f(p)).$$

Example 9.2.5
(1) $f : U \longrightarrow F$, $f(x) \equiv c$, $c \in F$.

$$f(x+h) - f(x) = 0_E = o\left(\|h\|_E\right).$$

So f is differentiable at every point of U and $\mathrm{D}(f(x)) = 0_F$.
(2) $\varphi \in \mathscr{L}(E, F)$.

$$\varphi(p+h) - \varphi(p) - \varphi(h) = 0_F = o\left(\|h\|_E\right).$$

So φ is differentiable at every point of E and $\mathrm{D}(\varphi(p)) = \varphi$.
(3) Let $(F_i, \|\cdot\|_i)$ be normed vector spaces over K, $i \in \{1, \ldots, n\}$, $n \in \mathbb{N}$. Suppose that $F = F_1 \oplus \cdots \oplus F_n$ and

$$\|\,(s_1, \ldots, s_n)\,\|_F = \max\{\|s_1\|_1, \ldots, \|s_n\|_n\}.$$

Let $U \subseteq E$ be an open subset, $f_i : U \longrightarrow F_i$ be a mapping.

$$f : U \longrightarrow F, \ f(x) = \left(f_1(x), \ldots, f_n(x)\right).$$

$$f(p+h) - f(p) = \left(f_1(p+h) - f_1(p), \ldots, f_n(p+h) - f_n(p)\right).$$

Suppose that each f_i is differentiable

$$\begin{aligned} & f(p+h) - f(p) - \left(\mathrm{D}f_1(p)(h), \ldots, \mathrm{D}f_n(p)(h)\right)|_F \\ = & \max_{i \in \{1, \ldots, n\}} \|f_i(p+h) - f_i(p) - \mathrm{D}f_i(p)(h)\|_{F_i} \\ = & o\left(\|h\|_E\right). \end{aligned}$$

So f is differentiable at p and

$$\mathrm{D}f(p)(h) = (\mathrm{D}f_1(p)(h), \ldots, \mathrm{D}f_n(p)(h)).$$

(4) Suppose that $E = K$. If $U \subseteq K$ is open and $f : U \longrightarrow F$ is differentiable at $p \in U$. We denote by $f'(p)$ the element $\mathrm{D}f(p)(1) \in F$.

$$f(p+h) - f(p) - \mathrm{D}f(p)(h) = o(\|h\|_E).$$

So

$$f(p+h) - f(p) - hf'(p) = o(\|h\|_E),$$
$$\frac{f(p+h) - f(p)}{h} - f'(p) = o(1).$$

That is,

$$\lim_{h\to 0} \frac{f(p+h) - f(p)}{h} = f'(p).$$

Theorem 9.2.6 Let $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$, $(G, \|\cdot\|_G)$ be normed vector spaces over a complete valued field $(K, |\cdot|)$. Let $U \subseteq E$ and $V \subseteq F$ be open subsets, $f : U \longrightarrow F$ and $g : V \longrightarrow G$ be mappings such that $f(U) \subseteq V$. Let $p \in U$. If f is differentiable at p and g is differentiable at $f(p)$, then $g \circ f : U \longrightarrow G$ is differentiable at p and

$$\mathrm{D}(g \circ f)(p)(h) = \mathrm{D}g(f(p))(\mathrm{D}f(p)(h)).$$

Proof

$$f(p+h) - f(p) - \mathrm{D}f(p)(h) = o(\|h\|_E),$$

so,

$$f(p+h) - f(p) = \mathcal{O}(\|h\|_E).$$

$$\begin{aligned} &g(f(p+h)) - g(f(p)) - \mathrm{D}g(f(p))(f(p+h) - f(p)) \\ =&o(\|f(p+h) - f(p)\|_F) = o(\mathcal{O}\|h\|_E) = o(\|h\|_E). \end{aligned}$$

$$\begin{aligned} &\mathrm{D}g(f(p))(f(p+h) - f(p)) - \mathrm{D}g(f(p))(\mathrm{D}f(p)(h)) \\ =&\mathrm{D}g(f(p))(f(p+h) - f(p) - \mathrm{D}f(p)(h)) \\ =&\mathcal{O}(o(\|h\|_E)) = o(\|h\|_E). \end{aligned}$$

So,

$$g(f(p+h)) - g(f(p)) - \mathrm{D}g(f(p))(\mathrm{D}f(p)(h)) = o(\|h\|_E).$$

$\square$

Remark 9.2.7 If $(E, \|\cdot\|_E) = (K, |\cdot|)$,

$$(g \circ f)'(p) = \mathrm{D}g(f(p))(f'(p)).$$

If $E = F = K$, $\|\cdot\|_E = \|\cdot\|_F = |\cdot|$.

$$(g \circ f)'(p) = g'(f(p)) \cdot f'(p).$$

Remark 9.2.8 Let $U \subseteq E$ be open. $f : U \longrightarrow F_1 \times \cdots \times F_n$. If f is differentiable at $p \in U$, for any $i \in \{1, \ldots, n\}$, the mapping

$$f_i := \pi_i \circ f : U \longrightarrow F_i$$

is differentiable at p and

$$\mathrm{D}(f_i)(p)(h) = \mathrm{D}\pi_i(f(p))\,(\mathrm{D}f(p)(h)) = \pi_i\,(\mathrm{D}f(p)(h))\,.$$

9.3 Multilineal Mappings

Definition 9.3.1 Let K be a commutative unitary ring. Let $E_1, \ldots, E_n; F$ be K-modules. We say that

$$\varphi : E_1 \times \ldots \times E_n \longrightarrow F$$

is n-linear if for any $i \in \{1, \ldots, n\}$ and any $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in E_1 \times \ldots \times E_{i-1} \times E_{i+1} \times \ldots \times E_n$, the mapping

$$E_i \longrightarrow F,\ x_i \mapsto \varphi(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)$$

is a homomorphism of K-modules. (K-linear mapping)
If $n = 1$, 1-linear is also called linear.
If $n = 2$, 2-linear is also called bilinear.

Example 9.3.2
(1) $K \times K \longrightarrow K\ (a, b) \longmapsto ab$ is bilinear.
(2) $K^n \times K^n \longrightarrow K\ (x, y) \longmapsto x \cdot y = \sum_{i=1}^{n} x_i y_i$ is bilinear.
(3) $K \times \ldots \times K \longrightarrow K\ (x_1, \ldots, x_n) \longmapsto x_1 \cdots x_n$ is n-linear.

Definition 9.3.3 We denote by $\mathrm{Hom}_K^{(n)}(E_1 \times \ldots \times E_n, F)$ the set of n-linear mappings from $E_1 \times \ldots \times E_n$ to F.

Definition 9.3.4 Let $(K, |\cdot|)$ be a complete valued field.
Let $(E_1, \|\cdot\|_{E_1}), \ldots, (E_n, \|\cdot\|_{E_n}), (F, \|\cdot\|_F)$ be normed vector spaces over K. For any $\varphi \in \mathrm{Hom}_K^{(n)}(E_1 \times \ldots \times E_n, F)$, we define

$$\|\varphi\| := \sup_{\substack{x_i \in E_i \setminus \{0\} \\ i \in \{1,\ldots,n\}}} \frac{\|\varphi(x_1, \ldots, x_n)\|_F}{\|x_1\|_{E_1} \cdots \|x_n\|_{E_n}}.$$

We denote by $\mathscr{L}(E_1 \times \cdots \times E_n, F)$ the set

$$\{\varphi \in \mathrm{Hom}_K^{(n)}(E_1 \times \cdots \times E_n, F) \mid \|\varphi\| < +\infty\}.$$

$\mathscr{L}^{(n)}(E_1 \times \ldots \times E_n, F)$ is a normed vector space of $\mathrm{Hom}_K^{(n)}(E_1 \times \ldots \times E_n, F)$, and the norm is $\|\cdot\|$.

Theorem 9.3.5 Let $(E_1, \|\cdot\|_{E_1}), \ldots, (E_n, \|\cdot\|_{E_n}), (F, \|\cdot\|_F)$ be normed vector spaces over K. Let $\varphi \in \mathscr{L}^{(n)}(E_1 \times \ldots \times E_n, F)$. For any $p = (p_1, \ldots, p_n) \in E_1 \times \ldots \times E_n$, φ is differentiable at p and

$$\mathrm{D}\varphi(p)(h_1, \ldots, h_n) = \sum_{i=1}^{n} \varphi(p_1, \ldots, p_{i-1}, h_i, p_{i+1}, \ldots, p_n).$$

Proof

$$\begin{aligned}\varphi(p+h) - \varphi(p) = \sum_{i=1}^{n} &\varphi(p_1 + h_1, \ldots, p_{i-1} + h_{i-1}, p_i + h_i, p_{i+1}, \ldots, p_n) \\ &-\varphi(p_1 + h_1, \ldots, p_{i-1} + h_{i-1}, p_i, p_{i+1}, \ldots, p_n)\end{aligned}$$

$$\begin{aligned}&\varphi(p_1 + h_1, \ldots, p_{i-1} + h_{i-1}, h_i, p_{i+1}, \ldots, p_n) \\ &-\varphi(p_1, \ldots, p_{i-1}, h_i, p_{i+1}, \ldots, p_n) \\ &= \sum_{j=1}^{i-1} \varphi(p_1 + h_1, \ldots, p_{j-1} + h_{j-1}, h_j, p_{j+1}, \ldots, h_i, \ldots, p_n).\end{aligned}$$

$$\begin{aligned}&\|\varphi(p_1+h_1,\ldots,p_{j-1}+h_{j-1},h_j,p_{j+1},\ldots,h_i,\ldots,p_n)\|_F\\ \leq&\|\varphi\|\cdot\prod_{k=1}^{j-1}\|p_k+h_k\|_{E_k}\cdot\|h_j\|_{E_j}\cdot\prod_{k=j+1}^{i-1}\|p_k\|_{E_k}\cdot\|h_i\|_{E_i}\cdot\prod_{k=i+1}^{n}\|p_k\|_{E_k}\\ =&\mathcal{O}\left(\|h\|^2\right)=o(h),\ h\to 0.\end{aligned}$$

□

Definition 9.3.6 Let K be a commutative unitary ring. $n\in\mathbb{N}_{\geq 1}$, E and F be K-modules. We say that

$$\varphi\in\mathrm{Hom}_K^{(n)}(E_1\times\ldots\times E_n,F)$$

is **symmetric** if

$$\forall\sigma\in\mathfrak{S}_{\{1,\ldots,n\}},\ \forall(x_1,\ldots,x_n)\in E^n,\ \varphi(x_{\sigma(1)},\ldots,x_{\sigma(n)})=\varphi(x_1,\ldots,x_n).$$

Let $P:E\longrightarrow F$ be a mapping. If there exists a symmetric $\varphi\in\mathrm{Hom}_K^{(n)}(E\times\ldots\times E,F)$ such that

$$\forall x\in E,\ P(x)=\varphi(x,\ldots,x),$$

we say that P is a **homogeneous polynomial mapping of degree** n.
If $F=K$, P is called a **homogeneous polynomial** on E. The symmetric polynomial mapping φ is called the **polarization** of P.

Proposition 9.3.7 Let $(K,|\cdot|)$ be a complete valued field that is non-trivial. Let $(E,\|\cdot\|_E)$ and $(F,\|\cdot\|_F)$ be normed vector spaces over K. Assume that $P:E\longrightarrow F$ is a homogeneous polynomial mapping of degree n. Which admits a bounded polarization φ. Then P is differentiable on E and,

$$\forall(x,h)\in E\times E,\ \mathrm{D}P(x)(h)=n\varphi(x,\ldots,x,h).$$

Proof Let

$$\begin{array}{rccl}\Delta: & E & \longrightarrow & E^n,\\ & x & \longmapsto & (x,\ldots,x).\end{array}$$

Then $P=\varphi\circ\Delta$. Since φ and Δ are differentiable, so it is P.

Moreover,

$$\begin{aligned}
\mathrm{D}P(x)(h) &= \mathrm{D}\varphi(\Delta(x))\,(\mathrm{D}\Delta(x)(h))\\
&= \mathrm{D}\varphi(x,\dots,x)(h,\dots,h)\\
&= \sum_{i=1}^{n} \varphi(x,\dots,x,h,x,\dots,x)\\
&= n\varphi(x,\dots,x,h).
\end{aligned}$$

□

Remark 9.3.8 Assume that $E = K$. Let $P : K \longrightarrow F$ be a homogeneous polynomial mapping of degree n of form $P(x) = x^n s$, where $s \in F$. Its polarization is of the form

$$\varphi(a_1,\dots,a_n) = a_1 \cdots a_n s.$$

$$P'(x) = \mathrm{D}P(x)(1) = n\varphi(x,\dots,x,1) = nx^{n-1}s.$$

Proposition 9.3.9 Let n be a positive integer $n \geq 2$. Let $(E_1, \|\cdot\|_1), \dots, (E_n, \|\cdot\|_n), (F, \|\cdot\|_F)$ be normed vector spaces. For any $i \in \{1,\dots,n_1\}$, the mapping

$$\mathscr{L}^{(n)}(E_1,\dots,E_n;F) \xrightarrow{f} \mathscr{L}^{(i)}(E_1,\dots,E_i,\mathscr{L}^{(n-i)}(E_{i+1},\dots,E_n;F))$$

$$\varphi \longmapsto \left(\substack{E_1\times\ldots e_n\longrightarrow\mathscr{L}^{(i)}(E_{i+1},\ldots,E_n;F)\\ (x_1,\ldots,x_i)\longmapsto\left(\substack{(x_{i+1},\ldots,x_n)\longmapsto\varphi(x_1,\ldots,x_n)\\ E_{i+1}\times\ldots\times E_n\in F}\right)}\right)$$

is an isomorphism of vector spaces over K, and in the same time an isometry, ($\|f(\varphi)\| = \|\varphi\|$).

Remark 9.3.10
$\forall (x_1,\dots,x_n) \in E_1 \times \ldots \times E_n$, $f(\varphi)(x_1,\dots,x_i)(x_{i+1},\dots,x_n) = \varphi(x_1,\dots,x_n)$

Proof $\forall (x_1,\dots,x_n) \in E_1 \times \ldots \times E_n$,

$$\begin{aligned}
\varphi(x_1,\dots,x_n) :&E_{i+1} \times \ldots \times E_n \longrightarrow F \text{ is bounded}\\
&(x_{i+1},\dots,x_n) \longmapsto \varphi(x_1,\dots,x_n)
\end{aligned}$$

Since

$$\|\varphi(x_1,\dots,x_n)\|_F \leq (\|\varphi\|\cdot\|x_1\|\ldots\|x_i\|)\,\|x_{i+1}\|\ldots\|x_n\|.$$

$$\begin{aligned}
\|f(\varphi)\| &= \sup_{x_j \in E_j \setminus \{0\}, j=1,\ldots,i} \frac{\|\varphi(x_1, \ldots, x_i, \cdot)\|}{\|x_1\|_{E_1} \cdots \|x_n\|_{E_i}} \\
&= \sup_{x_j \in E_j \setminus \{0\}, j=1,\ldots,i} \sup_{x_k \in E_k \setminus \{0\}, k=i+1,\ldots,n} \frac{\|\varphi(x_1, \ldots, x_n)\|_F}{\|x_1\|_{E_1} \cdots \|x_n\|_{E_n}} \\
&= \|\varphi\|.
\end{aligned}$$

Hence f is injective. ($\ker(f) = \{0\}$)
For any $\psi \in \mathscr{L}^{(i)}(E_1, \ldots, E_i, \mathscr{L}^{(n-i)}(E_{i+1}, \ldots, E_n))$,

$$\begin{aligned}
\varphi: \quad E_1 \times \ldots \times E_n &\longrightarrow F \\
(x_1, \ldots, x_n) &\longmapsto \psi(x_1, \ldots, x_i)(x_{i+1}, \ldots, x_n)
\end{aligned}$$

belongs to $\mathscr{L}^{(n)}(E_1, \ldots, E_n; F)$ and $f(\varphi) = \psi$. So f is surjective. $\square$

Corollary 9.3.11 If $E_1, \ldots, E_n$ are all finite dimensional, then

$$\mathscr{L}^{(n)}(E_1, \ldots, E_n; F) = \mathrm{Hom}_K^{(n)}(E_1 \times \ldots \times E_n, F).$$

Proof If $n = 1$, $\mathscr{L}(E_1, F) = \mathrm{Hom}_K(E_1, F)$.

$$\begin{aligned}
\mathscr{L}^{(n)}(E_1, \ldots, E_n; F) &\cong \mathscr{L}(E_1, \mathscr{L}^{(n-1)}(E_2, \ldots, E_n; F)) \\
&= \mathrm{Hom}_K(E_1, \mathrm{Hom}_K^{(n-1)}(E_2 \times \ldots \times E_n, F)) \\
&\cong \mathrm{Hom}_K^{(n)}(E_1 \times \ldots \times E_n, F).
\end{aligned}$$

$\square$

Let $(K, |\cdot|)$ be a complete nontrivial valued field. Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over K.

Definition 9.3.12 Let $U \subseteq E$ be an open subset of E, $f : U \longrightarrow F$ be a mapping.
If f is continuous on U, we say that f is **of class** $\mathcal{C}^0$ and we denote by

$$\mathrm{D}^0 f$$

the mapping $f : U \longrightarrow F$. Denote by

$$\mathcal{C}^0(U, F)$$

the set of mappings from U to F.

$$U \xrightarrow{(f,g)} K \times K \xrightarrow{\times} K$$

$$p \longmapsto (f(p), g(p)) \longmapsto f(p) \times g(p)$$

Let $p \in U$. If f is differentiable on an open neighborhood V of p such that $V \subseteq U$. Then

$$\begin{aligned} \mathrm{D}f : V &\longrightarrow \mathscr{L}(E, F) \\ x &\longmapsto \mathrm{D}f(x) \end{aligned}$$

is a mapping. If $\mathrm{D}f$ is $(n-1)$-times differentiable at p, we say that f is **of class** $\mathcal{C}^n$ at p. If f is of class $\mathcal{C}^n$ at every point of U, we say that f is **n-times differentiable** at p. We denote by

$$\mathrm{D}^n f(p) \in \mathscr{L}^{(n)}(E, \dots, E, F)$$

the n-linear mapping that sends $(h_1, \dots, h_n) \in E^n$ to

$$\mathrm{D}^{n-1}(\mathrm{D}f)(p)(h_1, \dots, h_{n-1})(h_n) \in F.$$

Remark 9.3.13

$$\mathrm{D}^n f(p)(h_1, \dots, h_n) = \mathrm{D}^i(\mathrm{D}^{n-i} f)(p)(h_1, \dots, h_i)(h_{i+1}, \dots, h_n).$$

9.4 Convexity

Definition 9.4.1 Let E be a vector space over a field K. $S \subseteq E$ be a non-empty subset.
We call affine combination of elements of S any element of E of the form

$$a_1 s_1 + a_2 s_2 + \cdots + a_n s_n,$$

where $n \in \mathbb{N}_{\geq 1}$, $s_1, \ldots, s_n \in S$, $a_1, \ldots, a_n \in K$ such that

$$a_1 + a_2 + \cdots + a_n = 1.$$

We denote by $\mathrm{Aff}(S)$ the set of all affine combinations of elements of S. One has $S \subseteq \mathrm{Aff}(S)$. $\mathrm{Aff}(S)$ is called the affine hull of S.
If $S = \mathrm{Aff}(S)$, we say that S is an affine subspace of E.

Proposition 9.4.2
(1) If F is a vector subspace of E, $\forall p \in E$,

$$p + F = \{p + x \mid x \in F\}$$

is an affine subspace of E.
(2) If $A \subseteq E$ is an affine subspace of E. For any $p \in A$,

$$A - p := \{x - p \mid x \in A\}$$

is a vector subspace of E, which is not dependent on the choice of p. We call it the vector space **associated** with A.

Proof
(1) Let $(x_1, \ldots, x_n) \in F^n$, $(a_1, \ldots, a_n) \in K^n$, such that $\sum_{i=1}^n a_i = 1$. Then

$$\begin{aligned}\sum_{i=1}^n a_i(p + x_i) &= p \cdot \sum_{i=1}^n a_i + \sum_{i=1}^n a_i x_i \\ &= p + \sum_{i=1}^n a_i x_i \in p + F.\end{aligned}$$

(2) Let $(x_1, \ldots, x_n) \in A^n$, $(b_1, \ldots, b_n) \in K^n$.

$$\begin{aligned}\sum_{i=1}^{n} b_i(x_i - p) &= \sum_{i=1}^{n} b_i x_i - \left(\sum_{i=1}^{n} b_i\right) p \\ &= \left(\sum_{i=1}^{n} b_i x_i + \left(1 - \sum_{i=1}^{n} b_i\right) p\right) - p \\ &\in A - p.\end{aligned}$$

Let $q \in A$, $\forall x \in A$, $x - p = (x - q) + (q - p) \in A - q$. So $A - p \subseteq A - q$. By symmetry, $A - q \subseteq A - p$. Hence $A - p = A - q$. □

Example 9.4.3 Let A be an m by p matrix with coefficients in $\mathbb{R}$. Let $(b_1, \ldots, b_n) \in E^m$. Consider the linear equation

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

The solution set is

$$S := \{(x_1, \ldots, x_p) \in E^p \mid A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}\}.$$

Claim: S is an affine subspace of E^p.

Proof Let $\underline{x}^{(1)}, \ldots, \underline{x}^{(n)}$ be elements of S, where $\underline{x}^{(i)} = (x_1^{(i)}, \ldots, x_p^{(i)})$. Let $(a_1, \ldots, a_n) \in \mathbb{R}^n$, $\underline{x} = a_1 x^{(1)} + \cdots + a_n x^{(n)}$.

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = A \left(a_1 \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_p^{(1)} \end{pmatrix} + \cdots + a_n \begin{pmatrix} x_1^{(n)} \\ \vdots \\ x_p^{(n)} \end{pmatrix} \right).$$

$$a_1 A \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_p^{(1)} \end{pmatrix} + \ldots a_n A \begin{pmatrix} x_1^{(n)} \\ \vdots \\ x_p^{(n)} \end{pmatrix} = (a_1 + \cdots + a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

$$x_j = a_1 x_j^{(1)} + \cdots + a_n x_j^{(n)}.$$

□

Proposition 9.4.4 Let $S \subseteq E$. Then $\mathrm{Aff}(S)$ is the smallest affine subspace of E containing S.

Proof
Let $A \subseteq E$ be an affine subspace containing S. $\forall n \in \mathbb{N}_{\geq 1}, \forall (x_1, \ldots, x_n) \in S^n \subseteq A^n$, $(a_1, \ldots, a_n) \in \mathbb{R}$, $a_1 + \cdots + a_n = 1$, one has

$$\sum_{i=1}^{n} a_i x_i \in A.$$

So $\mathrm{Aff}(S) \subseteq A$.

To show that $\mathrm{Aff}(S)$ is an affine subspace containing S, it is sufficient to check that $\mathrm{Aff}(S)$ is an affine subspace.

If $S = \varnothing$, then $\mathrm{Aff}(S) = \varnothing$. It is an affine subspace.
Suppose that $S \neq \varnothing$, $p \in S$. We prove that $\mathrm{Aff}(S) - p$ is equal to $\mathrm{Span}_{\mathbb{R}}(S - p)$. Let $y = a_1x_1 + \cdots + a_nx_n \in \mathrm{Aff}(S)$.

$$y - p = a_1(x_1 - p) + \cdots + a_n(x_n - p) \in \mathrm{Span}_{\mathbb{R}}(S - p).$$

Let $(x_1, \ldots, x_n) \in S^n$, $(b_1, \ldots, b_n) \in \mathbb{R}^n$.

$$\sum_{i=1}^{n} b_i(x_i - p) = \left(\sum_{i=1}^{n} b_i x_i + \left(1 - \sum_{i=1}^{n} b_i \right) p \right) - p \in \mathrm{Aff}(S) - p.$$

□

Definition 9.4.5 Let $S \subseteq E$. We call **convex combination** of elements of S any element of E of the form

$$a_1s_1 + a_2s_2 + \cdots + a_ns_n,$$

where $n \in \mathbb{N}_{\geq 1}$, $s_1, \ldots, s_n \in S$, $a_1, \ldots, a_n \in \mathbb{R}_{\geq 0}$ such that

$$a_1 + a_2 + \cdots + a_n = 1.$$

We denote by $\mathrm{Conv}(S)$ the set of all convex combinations of elements of S. $\mathrm{Conv}(S)$ is called the **convex hull** of S. One has $S \subseteq \mathrm{Conv}(S) \subseteq \mathrm{Aff}(S)$.

Proposition 9.4.6 Let E be a vector space over $\mathbb{R}$ and $C \subseteq E$. Then C is convex if and only if

$$\forall (x, y) \in C^2,\ \forall \lambda \in [0, 1],\ \lambda x + (1 - \lambda) y \in C.$$

Proof It is sufficient to check "$\Leftarrow$". We prove by induction on n that

$$\forall n \in \mathbb{N}_{\geq 1},\ \forall (x_1, \ldots, x_n) \in C^n,\ \forall (a_1, \ldots, a_n) \in \mathbb{R}^n_{\geq 0},\ \sum_{i=1}^{n} a_i = 1,\ \sum_{i=1}^{n} a_i x_i \in C.$$

The case where $n = 1$ is trivial. The case where $n = 2$ comes from the hypothesis. Suppose $n \geq 3$ in assuming that the statement holds for any integer less than n. If $a_n = 1$, then $a_1 = \cdots = a_{n-1} = 0$, so $\sum_{i=1}^{n} a_i x_i = x_n \in C$. If $a_n < 1$, we have $a_1 + \cdots + a_{n-1} = 1 - a_n > 0$. By the induction hypothesis,

$$x := \sum_{i=0}^{n-1} \frac{a_i}{1 - a_n} x_i \in C.$$

Taking $y = x_n$, $t = 1 - a_n$,

$$C \ni tx + (1 - t) y = \sum_{i=1}^{n} a_i x_i.$$

$\square$

9.5 Mean Value Theorems

Theorem 9.5.1 (Mean Value Inequality) Let $(F, \|\cdot\|_F)$ be normed vector spaces over $\mathbb{R}$. Let $(a, b) \in \mathbb{R}^2$ such that $a < b$. Let $f : [a, b] \longrightarrow F$ be a continuous mapping that is differentiable on $]a, b[$. Then

$$\|f(b) - f(a)\|_F \leq (b - a) \cdot \sup_{t \in]a,b[} \|f'(t)\|_F.$$

Proof We may suppose that $\sup_{t \in]a,b[} \|f'(t)\|_F < +\infty$. Take

$$M > \sup_{t \in]a,b[} \|f'(t)\|_F.$$

Let $m = \frac{a+b}{2}$. Let

$$J = \{x \in [m, b] \mid \forall t \in [m, x], \|f(t) - f(m)\|_F \leq M(t - m)\}.$$

It is an interval containing m. So it is of the form

$$[m, c[\text{ or } [m, c]$$

$$\forall t \in [m, c[, \ \|f(t) - f(m)\|_F \leq M(t - m).$$

Taking the limit $t < c$, $t \to c$, we get $c \in J$. So $J = [m, c]$. We then check $c = b$.
If $c \neq b$, then $c \in]a, b[$, so f is differentiable at c. That is

$$\|f(c+h) - f(c)\|_F = \|f'(c)h + o(\|h\|)\|_F \leq \|f'(c)\|_F h + o(\|h\|), \ h \to 0.$$

Since $M > \|f'(c)\|_F$, $\exists h_0 > 0$ such that

$$\forall h \in]0, h_0], \ \|f(c+h) - f(c)\|_F \leq Mh.$$

$$\begin{aligned}\|f(c+h) - f(m)\| &\leq \|f(c+h) - f(c)\| + \|f(c) - f(m)\| \\ &\leq Mh + M(c - m) = M(c + h - m).\end{aligned}$$

So $[m, c + h_0] \subseteq J$, contradiction. Thus $b = c$. $\|f(b) - f(m)\|_F \leq M(b - m)$.
By the same reason, $\|f(m) - f(a)\|_F \leq M(m - a)$. So

$$\|f(b) - f(a)\|_F \leq \|f(b) - f(m)\|_F + \|f(m) - f(a)\|_F \leq M(b - a).$$

Taking the limit when $M \to \sup\limits_{t \in]a,b[} \|f'(t)\|_F$, we get the announced result. □

Corollary 9.5.2 Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over $\mathbb{R}$. $U \subseteq E$ be an open subset, and $(x, y) \in U^2$ such that

$$[x, y] = \{tx + (1 - t)y \mid t \in [0, 1]\} \subseteq U.$$

Let $f : U \longrightarrow F$ be a differentiable mapping. Then

$$\|f(x) - f(y)\|_F \leq \left(\sup_{z \in]x,y[} \|\mathrm{D}f(z)\| \right) \cdot \|x - y\|_E.$$

Proof Let

$$\begin{aligned} g : \ [0,1] &\longrightarrow U \\ t &\longmapsto tx + (1-t)y. \end{aligned}$$

$$g(0) = x, \ g(1) = y, \ g'(t) = x - y.$$

Then,

$$(f \circ g)'(t) = \mathrm{D}f(g(t))(x-y),$$
$$\mathrm{D}(f \circ g)(t)(1) = \mathrm{D}f(g(t))(\mathrm{D}g(t)(1)).$$

By the theorem,

$$\begin{aligned}\|f(x) - f(y)\|_F &= \|f(g(1)) - f(g(0))\|_F \\ &\leq \sup_{t \in]0,1[} \|\mathrm{D}f(g(t))(x-y)\|_F \\ &\leq \sup_{t \in]0,1[} |\mathrm{D}f(g(t))| \cdot \|x-y\|_E \\ &= \sup_{z \in]x,y[} \|\mathrm{D}f(z)\| \cdot \|x-y\|_E.\end{aligned}$$

□

Definition 9.5.3 Let $(X, \mathscr{T})$ be a topological space, $p \in X$. Let U be a neighborhood of p and $f : U \longrightarrow \mathbb{R}$ be a mapping. If there exists a neighborhood V of p such that $p \in V \subseteq U$ and

$$\forall x \in V, \; f(p) \geq f(x),$$

we say that p is a **local maximum point** of f on U.

If p is a local maximum point or a local minimum point, we say that p is a **local extremum** of f on U.

If $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ are normed vector spaces. $U \subseteq E$ open, $f : U \to F$ is differentiable. If $p \in U$ is such that

$$\mathrm{D}f(p) = 0 \in \mathscr{L}(E, F),$$

we say that p is a **critical point** of f.

Theorem 9.5.4 Let $(E, \|\cdot\|)$ be a normed vector space over $\mathbb{R}$. $U \subseteq E$ be an open subset, $f : U \longrightarrow \mathbb{R}$ be a differentiable mapping. If $p \in U$ is a local extremum point of f, then it is a critical point $(\mathrm{D}f(p) = 0)$.

Proof There exists $r > 0$ such that $p + B(0, r) \subseteq U$ and

$$(h \in B(0, r)) \longmapsto f(p+h) - f(p) \in \mathbb{R}$$

does not change the sign.

$\forall h \in B(0,r),\ \forall \in [0,1],$

$$(f(p+th)-f(p))\,(f(p-th)-f(p)) \geq 0.$$

Taking the limit when $t \to 0$, $-\mathrm{D}f(p)(h)^2 \geq 0$. So $\mathrm{D}f(p)(h) = 0$. □

Theorem 9.5.5 (Rolle) Let $(a,b) \in \mathbb{R}^2$, $a < b$. Let $f : [a,b] \longrightarrow \mathbb{R}$ be a continuous mapping that is differentiable on $]a,b[$. If $f(a) = f(b)$, then

$$\exists t \in \,]a,b[,\ f'(t) = 0.$$

Proof If there exists t which is in $]a,b[$ and is an extremum point of f, then $f'(t) = 0$. Since $[a,b]$ is compact and f is continuous, so f attains its maximum and minimum.

If the extremum points of f are in $\{a,b\}$. Since $f(a) = f(b)$, f is compact, so $f'(t) = 0$ on $]a,b[$. □

Theorem 9.5.6 (Gronwall inequality) Let $(F, \|\cdot\|)$ be a normed vector space over $\mathbb{R}$, $(a,b) \in \mathbb{R}^2$, $a < b$. Let $f : [a,b] \longrightarrow F$ and $g : [a,b] \longrightarrow \mathbb{R}$ be differentiable mappings on $]a,b[$. If $\forall t \in \,]a,b[$, $\|f'(t)\| \leq g'(t)$, then

$$\|f(b)-f(a)\|_F \leq g(b)-g(a).$$

Proof Let $m \in \,]a,b[$. Let $\varepsilon > 0$,

$$J := \{t \in [m,b] \mid \forall s \in [m,t],\ \|f(s)-f(m)\|_F \leq g(s)-g(m)+\varepsilon(s-m)\}.$$

Since f and g are continuous, J is a closed interval of the form $[m,c]$.
If $c < b$,

$$\begin{aligned} f(c+h) &= f(c)+hf'(c)+o(h), \\ g(c+h) &= g(c)+hg'(c)+o(h), \quad h > 0,\ h \to 0. \end{aligned}$$

$\exists \delta > 0$, such that $[c, c+\delta] \subseteq [c,b]$ and $\forall h \in [0,\delta]$,

$$\|f(c+h)-f(c)\| \leq h\|f'(c)\| + \frac{\varepsilon}{2}h.$$

$$g(c+h)-g(c) \geq hg'(c) - \frac{\varepsilon}{2}h.$$

So,

$$\|f(c+h)-f(c)\| \leq g(c+h)-g(c)+\varepsilon h.$$

By the triangle inequality,

$$\|f(c+h)-f(m)\| \leq g(c+h)-g(m)+\varepsilon(c+h-m).$$

So $J \supseteq [m, c+\delta]$, contradiction.
Therefore $c=b$.

$$\|f(b)-f(m)\| \leq g(b)-g(m)+\varepsilon(b-m).$$

A similar argument shows that

$$\|f(m)-f(a)\| \leq g(m)-g(a)+\varepsilon(m-a).$$

Hence,

$$\|f(b)-f(a)\| \leq g(b)-g(a)+\varepsilon(b-a).$$

$$\|f(c+h)-f(c)+hf'(c)\| \leq \varphi(h)h,\ \lim_{h\to 0}\varphi(h)=0.$$

$$\exists\delta>0,\ \forall h>0,\ 0\leq h<\delta \Rightarrow |\varphi(h)| \leq \frac{\varepsilon}{2}.$$

□

Theorem 9.5.7 (Mean value theorem of Lagrange) Let $(a,b)\in\mathbb{R}^2$, $a<b$. Let $f:[a,b]\longrightarrow\mathbb{R}$ be a continuous mapping that is differentiable on $]a,b[$. Then

$$\exists\xi\in\,]a,b[,\ f(b)-f(a)=f'(\xi)(b-a).$$

Proof Let $g:[a,b]\longrightarrow\mathbb{R}$.

$$g(t):=f(b)-f(t)+C(b-t),\ \text{where } C=-\frac{f(b)-f(a)}{b-a}.$$

Then $g(a)=g(b)=0$, $g'(t)=-f'(t)-C$.

$$\exists\xi\in\,]a,b[,\ g'(\xi)=0,\ f'(\xi)=-C=\frac{f(b)-f(a)}{b-a}.$$

□

Theorem 9.5.8 (Darboux) Let I be an open interval in $\mathbb{R}$ and $f:I\longrightarrow\mathbb{R}$ be a differentiable mapping. Then $f'(I)$ is an interval.

Proof Let a, b be two elements in I such that $a < b$. Let

$$\begin{aligned} g:\ [a,b] &\longrightarrow \mathbb{R} \\ t &\longmapsto \begin{cases} \frac{f(t)-f(a)}{t-a}, & t \neq a \\ f'(a), & t = a \end{cases} \end{aligned}$$

g is continuous, and $g([a,b])$ is an interval. By the mean value theorem of Lagrange, $g([a,b]) \subseteq f'(I)$.
Let

$$\begin{aligned} h:\ [a,b] &\longrightarrow \mathbb{R} \\ t &\longmapsto \begin{cases} \frac{f(t)-f(b)}{t-b}, & t \neq b \\ f'(b), & t = b \end{cases} \end{aligned}$$

$h([a,b])$ is an interval contained in $f'(I)$.
$h([a,b]) \cup g([a,b])$ is an interval since

$$\frac{f(b)-f(a)}{b-a} \in h([a,b]) \cap g([a,b]),$$

$$\{f'(a), f'(b)\} \subseteq h([a,b]) \cup g([a,b]).$$

So the interval liking $f'(a)$, $f'(b)$ is contained in $f'(I)$. Hence, $f'(I)$ is an interval.
□

9.6 Higher Differential

We fix a complete non-trivially valued field $(K, |\cdot|)$. Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over K.

Definition 9.6.1 Let $U \subseteq E$ be an open subset, $f : U \longrightarrow F$ be a mapping, $p \in U$.
(1) If f is continuous at p, we say that f is 0-time differentiable at p, and we let

$$\mathrm{D}^0 f(p) := f(p).$$

(2) If f is differentiable at p, we say that f is 1-time differentiable at p, and we let

$$\mathrm{D}^1 f(p) := \mathrm{D}f(p).$$

(3) Let $n \geq 2$. If exists open neighborhood V of p such that $V \subseteq U$ and f is differentiable on V and $\mathrm{D}f$ is $n-1$-time differentiable on V, we say that f is

n-time differentiable at p, and we let

$$\mathrm{D}^n f(p) \in \mathscr{L}(E, \dots, E, F)$$

be the multilineal mapping sending $(h_1, \dots, h_n) \in E^n$ to

$$\mathrm{D}^{n-1}(\mathrm{D}f)(p)(h_1, \dots, h_{n-1})(h_n).$$

If $E = K$, $\mathrm{D}^n f(p)(1, \dots, 1)$ is denoted as $f^{(n)}(p) \in F$. $f^{(0)}(p)$ is often denoted as $f(p)$.

Remark 9.6.2 $\forall i \in \{1, \dots, n\}$,

$$\mathrm{D}^n f(p)(h_1, \dots, h_n) = \mathrm{D}^i(\mathrm{D}^{n-i} f)(p)(h_1, \dots, h_i)(h_{i+1}, \dots, h_n).$$

If $E = K$,

$$f^{(n)}(p)(h_1, \dots, h_n) = \mathrm{D}^{n-1}(\mathrm{D}f)(p)(h_1, \dots, h_{n-1})(h_n).$$

Definition 9.6.3 Let X be a set, we denote by $\mathfrak{S}_X$ the element of all bijection from X to X. $(\mathfrak{S}_X, \circ)$ forms a group. The identity mapping Id_X is the neutral element of $(\mathfrak{S}_X, \circ)$. $(\mathfrak{S}_X, \circ)$ is called the symmetric group of X. The elements of $(\mathfrak{S}_X, \circ)$ are called permutations of X.

Let $n \in \mathbb{N}_{\geq 2}$, $x_1, \dots, x_n$ be distinct elements of X. We denote by $(x_1\, x_2 \cdots x_n)$ the element of $\mathfrak{S}_X$ that sends x_i to x_{i+1}, $(i \in \{1, \dots, n-1\})$, x_n to x_1, $y \in X \backslash \{x_1, \dots, x_n\}$ to y itself. This element is called an n-cycle. A 2-cycle is also called a transposition.

Remark 9.6.4 $\mathfrak{S}_X$ acts on X.

$$\begin{array}{rcl} \mathfrak{S}_X \times X & \longrightarrow & X \\ (\sigma, x) & \longmapsto & \sigma(x). \end{array}$$

If $\sigma \in \mathfrak{S}_X$, $x \in X$, we denote by $\mathrm{orb}_\sigma(x)$ the set $\{\sigma^n(x) \mid n \in \mathbb{Z}\}$.

$$\langle \sigma \rangle := \{\sigma^n \mid n \in \mathbb{Z}\} \subseteq \mathfrak{S}_X$$

is a group. $\mathrm{orb}_\sigma(x)$ is the orbit of x under the action of $\langle \sigma \rangle$.

Proposition 9.6.5 If $\operatorname{orb}_\sigma(x)$ is finite of d elements, then $\sigma^d(x) = x$, and $\operatorname{orb}_\sigma(x) = \{x, \sigma(x), \dots, \sigma^{d-1}(x)\}$. Moreover, the restriction of σ to $\operatorname{orb}_\sigma(x)$ identifies to the restriction of the cycle $(x, \sigma(x), \dots, \sigma^{d-1}(x))$.

Proof Since $\operatorname{orb}_\sigma(x)$ is finite,

$$\{(n,m) \in \mathbb{Z}^2 \mid n < m,\ \sigma^n(x) = \sigma^m(x)\}$$

Let

$$l = \min\{m - n \mid (n,m) \in \mathbb{Z}^2\ n < m,\ \sigma^n(x) = \sigma^m(x)\}.$$

Then $x, \sigma(x), \dots, \sigma^{l-1}(x)$ are distinct, and $\sigma^l(x) = x$. $\forall n \in \mathbb{Z}$, then n can be written as $n = lp + r$, where $p \in \mathbb{Z}$, $r \in \{0, \dots, l-1\}$.

$$\sigma^n(x) = \sigma^r(\sigma^{lp}(x)) = \sigma^r((\sigma^l \circ \dots \sigma^l)(x)) = \sigma^r(x).$$

So, $\operatorname{orb}_\sigma(x) = \{x, \sigma(x), \dots, \sigma^{l-1}(x)\}$, ($l = d$). □

Remark 9.6.6 If X is finite, then X can be written as a distinct union of orbits (under the action of $\langle\sigma\rangle$). Let $d_i = \#(\operatorname{orb}_\sigma(x_i))$, $i = 1, \dots, n$, then

$$\sigma|_{\operatorname{orb}_\sigma(x^{(i)})} = \left(x^{(i)}, \sigma(x^{(i)}), \dots, \sigma^{d_i-1}(x^{(i)})\right)|_{\operatorname{orb}_\sigma(x^{(i)})}.$$

So $\sigma = \tau_1 \circ \cdots \circ \tau_n$, where $\tau_i := (x^{(i)}, \sigma(x^{(i)}), \dots, \sigma^{d_i-1}(x^{(i)}))$.

Corollary 9.6.7 Suppose that X is finite. Any $\sigma \in \mathfrak{S}_X$ can be written as a composition of transpositions.

Proof

$$(x_1\ \dots\ x_n) = (x_1\ x_2) \circ (x_2 \dots\ x_n),$$

So,

$$(x_1\ \dots\ x_n) = (x_1\ x_2) \circ (x_2\ x_3) \circ \cdots \circ (x_{n-1}\ x_n).$$

□

Definition 9.6.8 Denote by $\mathfrak{S}_n$ the symmetric group $\mathfrak{S}_{\{1,\dots,n\}}$. A composition of the form $(i\ i+1)$, $i \in \{1, \dots, n-1\}$ is called an adjacent transposition.

Corollary 9.6.9 Any $\sigma \in \mathfrak{S}_n$ can be written as a composition of adjacent transpositions.

Proof Let $(j,k) \in \{1,\ldots,n\}^2$, $j < k$,

$$(j-1\ j) \circ (j\ k) \circ (j-1\ j) = (j-1\ k).$$

$$(j\ k) = (j\ j+1) \circ (j+1\ j+2) \circ \cdots \circ (k-1\ k) \circ \ldots (j\ j+1).$$

□

Theorem 9.6.10 (Schwarz) Let $U \subseteq E$ be an open subset, $f : U \longrightarrow F$ be a mapping. $n \in \mathbb{N}_{\geq 1}$, $p \in U$. Assume that f is n-times differentiable at p. Then $\forall \sigma \in \mathfrak{S}_n$, $\forall (h_1, \ldots h_n) \in E^n$,

$$\mathrm{D}^n f(p)(h_1, \ldots, h_n) = \mathrm{D}^n f(p)(h_{\sigma(1)}, \ldots, h_{\sigma(n)}).$$

Proof (By induction) The case where $n = 1$ is trivial. Case $n = 2$: Exists V open, $p \in V \subseteq U$. f is differentiable on V and $\mathrm{D}f$ is differentiable at p.

$$\mathrm{D}f(p+h)(\cdot) - \mathrm{D}f(p)(\cdot) - \mathrm{D}^2 f(p)(h,\cdot) = o\left(\|h\|_E\right).$$

Let $\varepsilon > 0$, $\exists \delta > 0$, $\forall h \in E$, $\|h\|_E \leq 2\delta \Rightarrow p + h \in V$ and

$$\|\mathrm{D}f(p+h)(\cdot) - \mathrm{D}f(p)(\cdot) - \mathrm{D}^2 f(p)(h,\cdot)\| \leq \varepsilon \|h\|_E.$$

Let $h \in E$ such that $\|h\|_E \leq \delta$. Define $g_h : B(0,\delta) \longrightarrow F$ as

$$g_h(k) = f(p+h+k) - f(p+h) - f(p+k) + f(p) - \mathrm{D}^2 f(p)(h,k).$$

Then,

$$\begin{aligned}
\mathrm{D}g_h(k)(\cdot) =& \mathrm{D}f(p+h+k)(\cdot) - \mathrm{D}f(p+k)(\cdot) - \mathrm{D}^2 f(p)(h,\cdot) \\
=& \mathrm{D}f(p+h+k)(\cdot) - \mathrm{D}f(p)(\cdot) - \mathrm{D}^2 f(p)(h+k,\cdot) \\
& -(\mathrm{D}f(p+k)(\cdot) - \mathrm{D}f(p)(\cdot) - \mathrm{D}^2 f(p)(k,\cdot))
\end{aligned}$$

$$\|\mathrm{D}g_h(k)(\cdot)\| \leq \varepsilon \|h+k\|_E + \varepsilon \|k\|_E \leq 3\varepsilon \max\{\|k\|_E, \|h\|_E\}.$$

$g_h(0) = 0$. Therefore, $\|g_h(k)\| \leq 3\varepsilon \max \|k\|_E, \|h\|_E{}^2$ (mean value inequality).

$$\|g_h(k) - g_h(0)\| \leq \left(\sup_{t \in]0,1[} \|\mathrm{D}g_h(tk)\|\right) \cdot \|k\|.$$

Therefore,

$$f(p+h+k)-f(p+h)-f(p+k)+f(p)-\mathrm{D}^2 f(p)(h,k) = o\left(\max\{\|k\|_E, \|h\|_E\}^2\right).$$

By symmetry,

$$f(p+h+k)-f(p+h)-f(p+k)+f(p)-\mathrm{D}^2 f(p)(k,h) = o\left(\max\{\|k\|_E, \|h\|_E\}^2\right).$$

$$\mathrm{D}^2 f(p)(h,k) - \mathrm{D}^2 f(p)(k,h) = o\left(\max\{\|h\|_E, \|k\|_h\}^2\right).$$

$$\mathrm{D}^2 f(p)(th,tk) - \mathrm{D}^2 f(p)(tk,th) = o(|t|^2),\ t \to 0.$$

$$\mathrm{D}^2 f(p)(h,k) - \mathrm{D}^2 f(p)(k,h) = o(1),\ t \to 0.$$

Suppose $n \geq 3$.

$$\mathrm{D}^n f(p)(h_1,\ldots,h_n) = \mathrm{D}^{n-1}(\mathrm{D}f)(p)(h_1,\ldots,h_{n-1})(h_n).$$

If $\sigma = (j\ j+1)$, $j \leq 2$,

$$\mathrm{D}^{n-1}(\mathrm{D}f)(p)(h_1,\ldots,h_{n-1})(h_n) = \mathrm{D}^{n-1}(\mathrm{D}f)(p)(h_{\sigma(1)},\ldots,h_{\sigma(n-1)})(h_n)$$

by the induction hypothesis, if $\sigma = (n-1\ n)$,

$$\mathrm{D}^n f(p)(h_1,\ldots,h_n) = \mathrm{D}^2\left((\mathrm{D}^{n-2} f)(h_1,\ldots,h_{n-2})(h_{n-1},h_n)\right)$$

$$\begin{aligned}
\mathrm{D}^n f(p)(h_{\sigma(1)},\ldots,h_{\sigma(n)}) &= \mathrm{D}^n f(p)(h_1,\ldots,h_{n-1}) \\
&= \mathrm{D}^2\left((\mathrm{D}^{n-2} f)(h_1,\ldots,h_{n-2})(h_n,h_{n-1})\right) \\
&= \mathrm{D}^n f(p)(h_1,\ldots,h_n).
\end{aligned}$$

$\square$

9.7 Taylor's Formula

Theorem 9.7.1 (Toylor-Young) Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over $\mathbb{R}$, $U \subseteq E$ open, $n \in \mathbb{N}$, $f : U \longrightarrow F$ be a mapping, $p \in U$. Suppose that f is n-times differentiable at p. Then

$$f(x) = f(p) + \sum_{k=1}^{n} \frac{1}{k!} \mathrm{D}^k f(p)(x-p, \ldots, x-p) + o(\|x-p\|^n),\ x \to p.$$

Proof (By induction on n)
$n = 0$, $f(x) = f(p) + o(1)$ follows by continuity of f; $n = 1$ follows by the differentiability of f.

From $n-1$ to n. Let $g : U \longrightarrow F$

$$g(x) = f(x) - f(p) - \sum_{k=1}^{n} \frac{1}{k!} \mathrm{D}^k f(p)(x-p, \ldots, x-p).$$

g is differentiable on an open neighborhood of p,

$$\mathrm{D}g(x)(h) = \mathrm{D}f(x)(h) - \sum_{k=1}^{n} \frac{1}{k!} k \mathrm{D}^k f(p)(x-p, \ldots, x-p, h)$$

$$\mathrm{D}g(x) = \mathrm{D}f(x) - \sum_{l=0}^{n-1} \frac{1}{l!} \mathrm{D}^l(\mathrm{D}f)(x-p, \ldots, x-p) \overset{\text{hyp.}}{=\!=} o(\|x-p\|^{n-1}), x \to p.$$

So $g(x) = o(\|x-p\|^n)$.

$$\forall \varepsilon > 0,\ \exists \delta > 0, \forall x \in B(p, \delta),\ \|\mathrm{D}g(x)\| \leq \varepsilon \|x-p\|^{n-1}.$$

$g(p) = 0$, so

$$\|g(x) - g(p)\| \leq \varepsilon \|x-p\|^{n-1} \cdot \|x-p\| = \varepsilon \|x-p\|^n.$$

□

Theorem 9.7.2 (Taylor-Lagrange) Let $(a, b) \in \mathbb{R}^2$, $a < b$. $f : [a, b] \longrightarrow \mathbb{R}$ be a continuous mapping. Suppose that f is $(n+1)$-times differentiable on $]a, b[$ and $\forall k \in \{3, \ldots, n\}$, $f^{(k)} :]a, b[\longrightarrow \mathbb{R}$ tends to a continuous mapping $[a, b] \longrightarrow \mathbb{R}$.

Then

$$\exists \xi \in]a, b[, f(b) - \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) = \frac{f^{(n+1)}(\xi)(b-a)^{n+1}}{(n+1)!}.$$

Proof Let $g : [a, b] \longrightarrow \mathbb{R}$.

$$g(t) := \sum_{k=0}^{n} \frac{(b-t)^k}{k!} f^{(k)}(t) - C \frac{(b-t)^{n+1}}{(n+1)!}.$$

Then $g(b) = f(b)$, $g(a) = \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) - C \frac{(b-a)^{n+1}}{(n+1)!}$.

$$\begin{aligned} g'(t) &= \sum_{k=0}^{n} \frac{(b-t)^k}{k!} f^{(k+1)}(t) - \sum_{k=1}^{n} \frac{(b-t)^{k-1}}{(k-1)!} f^{(k)}(t) + C \frac{(b-t)^n}{n!} \\ &= \frac{(b-t)^n}{n!} f^{(n+1)}(t) + C \frac{(b-t)^n}{n!}. \end{aligned}$$

Take C such that $g(a) = g(b)$. Then by Rolle's theorem, $\exists \xi \in]a, b[, g'(\xi) = 0$, $C = -f^{(n+1)}(\xi)$. Then,

$$g(a) = \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{f^{(n+1)}(\xi)}{(n+1)!} + \frac{f^{(n+1)}(\xi)}{(n+1)!}(b-a)^{n+1} = f(b) = g(b).$$

□

Theorem 9.7.3 Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be two normed vector spaces over $\mathbb{R}$, $U \subseteq E$ be an open subset, and $f : U \longrightarrow F$ be a mapping that is $(n+1)$-times differentiable, where $n \in \mathbb{N}$. Let $p \in U$, $h \in E$ such that $\forall t \in [0, 1]$, $p + th \in U$. Let

$$M = \sup_{t \in [0,1]} \|\mathrm{D}^{n+1} f(p + th)\|.$$

Then,

$$\|f(p+h) - \sum_{k=0}^{n} \frac{1}{k!} \mathrm{D}^k f(p)(h, \dots, h)\|_F \leq \frac{M}{(n+1)!} \|h\|_E^{n+1}.$$

Proof We define $\phi : [0,1] \longrightarrow F$

$$\phi(t) = f(p+th) + \sum_{k=1}^{n} \frac{(1-t)^k}{k!} \mathrm{D}^k f(p+th)(h, \ldots, h).$$

$$\phi(0) = \sum_{k=0}^{n} \frac{1}{k!} \mathrm{D}^k f(p)(h, \ldots, h), \ \phi(1) = f(p+h).$$

$$\begin{aligned}
\phi'(t) =& \mathrm{D} f(p+h)(h) + \sum_{k=1}^{n} \frac{(1-t)^k}{k!} \mathrm{D}^{k+1} f(p+th)(h, \ldots, h) \\
& - \sum_{k=1}^{n} \frac{(1-t)^{k-1}}{(k-1)!} \mathrm{D}^k f(p+th)(h, \ldots, h) \\
& = \sum_{k=0}^{n} \frac{(1-t)^k}{k!} \mathrm{D}^{k+1} f(p+th)(h, \ldots, h) \\
& - \sum_{l=0}^{n-1} \frac{(1-t)^l}{(l)!} \mathrm{D}^{l+1} f(p+th)(h, \ldots, h) \\
& = \frac{(1-t)^n}{n!} \mathrm{D}^{n+1} f(p+th)(h, \ldots, h).
\end{aligned}$$

So,

$$\|\phi'(t)\| \leq M \|h\|_E^{n+1} \frac{(1-t)^n}{n!}, \ t \in [0,1].$$

By Gronwall's inequality,

$$\|\phi(1) - \phi(0)\|_F \leq M \cdot \|h\|^{n+1} \frac{1}{(n+1)!}.$$

□

9.8 Banach Space

Proposition 9.8.1 Let (X, d) be a metric space and $(x_n)_{n \in \mathbb{N}}$ be a sequence in X. If

$$\sum_{n \in \mathbb{N}} \mathrm{d}(x_n, x_{n+1}) < +\infty,$$

then $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence.

Proof Let $N \in \mathbb{N}$. If $(n,m) \in \mathbb{N}^2_{\geq N}$, $n > m$, by the triangle inequality,

$$\mathrm{d}(x_m, x_n) \leq \sum_{k=m}^{n-1} \mathrm{d}(x_k, x_{k+1}) \leq \sum_{k\geq N} \mathrm{d}(x_k, x_{k+1}).$$

So,

$$0 \leq \sup_{(n,m)\in\mathbb{N}_{\geq N}} \mathrm{d}(x_n, x_m) \leq \sum_{k\geq N} \mathrm{d}(x_k, x_{k+1}).$$

Taking the limit when $N \to +\infty$, we get

$$\lim_{N\to+\infty} \sup_{(n,m)\in\mathbb{N}^2_{\geq N}} \mathrm{d}(x_n, x_m) = 0.$$

Let $(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. If $(\sum_{k=0}^{n} a_k)_{n\in\mathbb{N}}$ converges to some l in $\mathbb{R}$. Then, $l - \sum_{k=0}^{N-1} a_k$ converges to 0. If $a_k \leq 0$ for any $k \in \mathbb{N}$, $l - \sum\limits_{k=0}^{N-1} a_k = \sum\limits_{k=N}^{+\infty} a_k$.

$$l - \sum_{k=0}^{N-1} a_k = \lim_{n\to+\infty} \left(\sum_{k=0}^{n} a_k - \sum_{k=0}^{N-1} a_k \right) = \lim_{n\to+\infty} \sum_{k=N}^{n} a_k.$$

$\square$

Definition 9.8.2 Let $(K, |\cdot|)$ be a complete valued field and $(E, \|\cdot\|)$ be a normed vector space over K. If E equipped with the metric

$$\begin{array}{rcl} E \times E & \longrightarrow & \mathbb{R}_{\geq 0} \\ (x,y) & \longmapsto & \|x-y\|_E. \end{array}$$

is complete, we say that $(E, \|\cdot\|)$ is a **Banach space**.
Let $(E, \|\cdot\|)$ be a Banach space. If $(x_n)_{n\in\mathbb{N}}$ is a sequence in E such that $\sum_{n\in\mathbb{N}} \|x_n\| < +\infty$, we say that $\sum_{n\in\mathbb{N}} x_n$ **converges absolutely**.

Remark 9.8.3 Suppose that $\sum\limits_{n\in\mathbb{N}} x_n$ converges absolutely. Then $\left(\sum\limits_{k=0}^{n} x_k\right)_{n\in\mathbb{N}}$ is a Cauchy sequence, since

$$\|x_n\| = \|\sum_{k=0}^{n} x_k - \sum_{k=0}^{n} x_{k-1}\|.$$

So, $\sum\limits_{n\in\mathbb{N}} x_n$ converges.

Theorem 9.8.4 (Root test of Cauchy) Let $(E, \|\cdot\|)$ be a Banach space and $(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. Let

$$r = \limsup_{n\to+\infty} \|x_n\|^{\frac{1}{n}} \in [0, +\infty]$$

If $r < 1$, then $\sum_{n\in\mathbb{N}} x_n$ converges absolutely.
If $r > 1$, then $\sum_{n\in\mathbb{N}} x_n$ diverges.

Lemma 9.8.5 If a series $\sum_{n\in\mathbb{N}} x_n$ converges, then $\lim_{n\to+\infty} \|x_n\| = 0$.

Proof (of lemma)

$$\|x_n\| = \|\sum_{k=0}^{n} x_k - \sum_{k=0}^{n-1} x_k\|.$$

Since $\sum_{k}^{n} x_k$ converges to some $l \in E$.

$$\lim_{n\to+\infty} \|x_n\| = \lim_{n\to+\infty} \|\sum_{k=0}^{n} x_k - \sum_{k=0}^{n-1} x_k\| = \|l - l\| = 0.$$

□

Proof (of theorem) If $r > 1$, $\exists \beta > 1$ such that $r > \beta$. Since $r = \limsup_{n\to+\infty} \|x_n\|^{\frac{1}{n}}$, $\exists I \subseteq \mathbb{N}$ infinite such that $\lim_{n\in I, n\to+\infty} \|x_n\|^{\frac{1}{n}} = r$ (Bolzano-Weierstrass).

$$\exists N \in \mathbb{N}, \forall n \in I_{\geq N}, \|x_n\|^{\frac{1}{n}} \geq \beta.$$

So, $\|x_n\| \geq \beta^n \geq 1$. So $\sum_{n\in\mathbb{N}} x_n$ diverges.
If $r < 1$, $\exists \alpha \in]0, 1[, r < \alpha$. Since $r = \limsup_{n\to+\infty} \|x_n\|^{\frac{1}{n}}$,

$$\exists N \in \mathbb{N},\ \forall n \geq N,\ \|x_n\|^{\frac{1}{n}} \leq \alpha, \|x_n\| \leq \alpha^n.$$

So,

$$\sum_{n\geq N} \|x_n\| \leq \sum_{n\geq N} \alpha^n = \frac{\alpha^N}{1-\alpha} < +\infty.$$

Therefore, $\sum_{n\in\mathbb{N}} x_n$ converges absolutely. □

Theorem 9.8.6 (Ratio test of D'Alembert) Let $(E, \|\cdot\|)$ be a Banach space and $(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$.
(1) If

$$\limsup_{n\to+\infty} \frac{\|x_{n+1}\|}{\|x_n\|} < 1,$$

then $\sum_{n\in\mathbb{N}} x_n$ converges absolutely.
(2) If

$$\limsup_{n\to+\infty} \frac{\|x_{n+1}\|}{\|x_n\|} > 1,$$

then $\sum_{n\in\mathbb{N}} x_n$ diverges.

Proof
(1) Let $0 < \alpha < 1$ such that

$$\limsup_{n\to+\infty} \frac{\|x_{n+1}\|}{\|x_n\|} < \alpha.$$

$$\exists N \in \mathbb{N}, \forall x \in \mathbb{N}_{\geq N}, \|x_{n+1}\| \leq \alpha\|x_n\| \leq \alpha^{n+1-N}\|x_N\|.$$

Thus,

$$\sum_{n\geq\mathbb{N}} \|x_n\| \leq \sum_{n\geq\mathbb{N}} \|x_N\|\alpha^{n-N} = \|x_N\|\frac{1}{1-\alpha} < +\infty.$$

(2) Let $\beta > 1$ such that

$$\liminf_{n\to+\infty} \frac{\|x_{n+1}\|}{\|x_n\|} > \beta.$$

$$\exists N \in \mathbb{N}, x_N \neq 0, \text{ and } \forall n \in \mathbb{N}_{\geq N}, \|x_{n+1}\| \geq \beta\|x_n\|$$

$$\forall n \geq N, \|x_n\| \geq \beta^{n-N}\|x_N\| \to +\infty \; (n \to +\infty)$$

So $\sum_{n\in\mathbb{N}} x_n$ diverges. □

Let $z \in \mathbb{C}$. The series $\sum_{n\in\mathbb{N}} \frac{z^n}{n!}$ converges absolutely since

$$\left|\frac{z^{n+1}/(n+1)!}{z^n/n!}\right| = \frac{|z|}{n+1} \to 0 \; (n \to +\infty).$$

We denote by e^z this limit.

9.9 Local inversion

Definition 9.9.1 Let X be a topological space and $Y \subseteq X$. If $\overline{Y} = X$, we say that Y is dense.

Theorem 9.9.2 (Baire) Let (X, d) be a complete metric space. Let $(\Omega_n)_{n\in\mathbb{N}}$ be a sequence of dense open subset of X. Let $\Omega = \bigcap_{n\in\mathbb{N}} \Omega_n$, then Ω is dense in X.

Proof Suppose that Ω is not dense. Let $x_0 \in X\backslash\overline{\Omega}$, exists $\varepsilon > 0$ such that $B(x_0, \varepsilon) \subseteq X\backslash\overline{\Omega}$.
Let $r_0 = \varepsilon$. We construct in a recursive way sequence $(x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}}$ and $(r_n)_{n\in\mathbb{N}} \in \mathbb{R}^n_{\geq 0}$ as follows.

Suppose that (x_n, r_n) is chosen. $B(x_n, r_n) \cap \Omega_n \neq \varnothing$. We pick $x_{n+1} \in X$ and $r_{n+1} \leq \frac{x_n}{2}$ such that $B(x_{n+1}, r_{n+1}) \subseteq B(x_n, r_n) \cap \Omega_n$, $\mathrm{d}(x_{n+1}, x_n) < r_n$. $\sum_{n\in\mathbb{N}} r_n < +\infty$ (ratio test).

Then the sequence converges to some l. For any $n \in \mathbb{N}$, $x_n \in B(x_0, \varepsilon)$. So $l \in \overline{B}(x_0, \varepsilon)$.

Moreover, $\forall n \in \mathbb{N}$, $l \in \overline{B}(x_{n+1}, r_{n+1}) \subseteq B_{x_n, r_n} \cap \Omega_n$. Thus $l \in \cap_{n\in\mathbb{N}}\Omega_n = \Omega$. Contradiction. □

Corollary 9.9.3 Let (X, d) be a non-empty complete metric space and $(Y_n)_{n\in\mathbb{N}}$ be a family of closed subsets of X such that $X = \bigcup_{n\in\mathbb{N}} Y_n$. Then exists $n \in \mathbb{N}$ such that $Y^\circ \neq \varnothing$.

Proof Let $\Omega_n = X\backslash Y_n$. Suppose that $\forall n \in \mathbb{N}$, $Y_n^\circ = \varnothing$. Then $\overline{\Omega}_n = X\backslash Y_n^\circ = X$. Thus $\Omega := \bigcap_{n\in\mathbb{N}} \Omega_n$ is dense in X. Namely, $X = \overline{\Omega}$. So

$$\varnothing = X\backslash\overline{\Omega} = (X\backslash\Omega)^\circ = \left(X\backslash \bigcap_{n\in\mathbb{N}} \Omega_n\right)^\circ = \left(\bigcup_{n\in\mathbb{N}} Y_n\right)^\circ = X^\circ = X.$$

Contradiction. □

Theorem 9.9.4 (Banach) Let $(K, |\cdot|)$ be a complete non-trivially valued filed, and E be a vector space over K. Let $\|\cdot\|_1$, $\|\cdot\|_2$ be two norms on E such that $(E, \|\cdot\|_1)$ and $(E, \|\cdot\|_2)$ are both Banach spaces.

If $\exists C > 0$ such that $\|\cdot\|_2 \leq C\|\cdot\|_1$. Then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent. $(\exists C' > 0, \|\cdot\|_1 \leq C'\|\cdot\|_2)$

Proof For $x \in E$ and $r > 0$. Let

$$B_i(x,r) := \{y \in E \mid \|y-x\|_i < r\},\ i = 1,2$$

$\forall y \subseteq E$, let $\overline{Y}^{\|\cdot\|_2}$ be the closure of Y in $(E, \|\cdot\|_2)$.

$$E = \bigcup_{n\geq 1} B_1(0,n) = \bigcup_{n \geq 1} \overline{B_1(0,n)}^{\|\cdot\|_2}.$$

Hence, $\exists n_0 \geq 1, p \in E, r_0 > 0$ such that

$$B_2(p, r_0) \subseteq \overline{B_1(0,n_0)}^{\|\cdot\|_2}$$

or equivalently,

$$B_2(0,r_0) \subseteq \overline{B_1(-p,n_0)}^{\|\cdot\|_2} \subseteq \overline{B_1(0, n_0 + \|p\|_1)}^{\|\cdot\|_2}$$

since $\forall x \in B_1(-p, n_0)$

$$\|x\|_1 = \|x-p+p\|_1 \leq \|x-p\| + \|p\|_1 < n_0 + \|p\|_1.$$

Let $r_1 = n_0 + \|p\|_1$,

$$B_2(0,r_0) \subseteq \overline{B_1(0,r_1)}^{\|\cdot\|_2} \subseteq B_1(0,r_1) + B_2(0, r_0|a|)$$

where $a \in K, 0 < |a| < \frac{1}{2}$.
In fact, $\forall x \in \overline{B_1(0,r_0)}^{\|\cdot\|_2}$, exists sequence $(x_n)_{n\in\mathbb{N}} \in B_1(0,r_1)^{\mathbb{N}}$, such that $x_n \to x$ in $(E, \|\cdot\|_2)$, $\exists n \in \mathbb{N}$, $\|x_n - x\|_2 < r_0|a|$

$$B_2(0, r_0|a|^n) \subseteq B_1(0, r_1|a|^n) + B_2(0, r_0|a|^{n+1})$$

Let $y \in B_2(0,r_0)$, we choose $(x_0, y_0) \in B_1(0,r_1) \times B_2(0, r_0|a|)$ such that $y = x_0 + y_0$. When (x_n, y_n) si chosen, let $(x_{n+1}, y_[n+1]) \in B_1(0, r_0|a|^{n+1}) \times B_2(0, r_0|a|^{n+2})$, $y_n = x_{n+1} + y_{n+1}$, $y = y_n + \sum_{k=0}^n x_k$. So $\sum_{n\in\mathbb{N}} x_n$ converges to y.
Moreover, $\sum_{n\in\mathbb{N}} \|x_n\| < +\infty$, so it converges in $(E, \|\cdot\|_1)$ to some x. Therefore, $x = y$ since $\|\cdot\|_2 \leq C\|\cdot\|_1$. So $\|y\|_1 \|x\|_1 \leq \sum_{n\in\mathbb{N}} \|x_n\|_1 \leq \frac{r_1}{1-|a|}$.

Therefore $B_2(0,r_0) \subseteq B_1(0, \frac{r}{1-|a|})$. So $\|\cdot\|_1$ is bounded by a constant times $\|\cdot\|_2$. □

Proposition 9.9.5 Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be Banach spaces over a complete non-trivially valued filed $(K, |\cdot|)$, and $f : E \longrightarrow F$ be a bounded mapping.
(1) If f is invertible, then f^{-1} is bounded.
(2) If f is surjective, for any $U \subseteq E$ open, $f(U)$ is open in F.

Proof
(1) We define a mapping

$$\begin{array}{rccc} \|\cdot\|'_E : & E & \longrightarrow & \mathbb{R}_{\geq 0} \\ & x & \longmapsto & \|f(x)\|_F. \end{array}$$

This is a norm on E. In fact, if $\|x\|'_E = \|f(x)\|_F = 0$, then $f(x) = 0_F$. So $x = 0_E$. Moreover,

$$\forall x \in E,\ \|x\|'_E = \|f(x)\|_F \leq \|f\|\|x\|_E.$$

So there exists $C > 0$ such that $\|\cdot\|_E \leq C\|\cdot\|'_E$. That is,

$$\forall y \in F,\ \|y\|_F = \|f(f^{-1}(y))\|_F = \|f^{-1}(y)\|'_E \geq C^{-1}\|f^{-1}(y)\|_E.$$

So, $\|f^{-1}\| \leq C$.
(2) Let

$$E_0 = \ker(f) = \{x \in E \mid f(x) = 0_F\}.$$

This is a closed vector subspace of E. $\|\cdot\|_E$ induces by passing to quotient a norm $\|\cdot\|_Q$ on $Q := E/E_0$. Let

$$\begin{array}{rccc} g : & Q & \longrightarrow & F \\ & [x] & \longmapsto & f(x). \end{array}$$

This is a K-linear bijection.

If $\alpha \in Q$,

$$\forall x \in \alpha,\ \|g(\alpha)\|_F = \|f(x)\|_F \leq \|f\| \cdot \|x\|_E.$$

Since $\|\alpha\|_Q := \inf_{x\in\alpha} \|x\|_E$, $\|g(\alpha)\|_F \leq \|f\| \cdot \|\alpha\|_Q$. So $\|g\| \leq \|f\|$. By (1), g^{-1} is bounded (hence is continuous).

If $V \subseteq Q$ is open, then $g(V) \subseteq F$ is open. Let $U \subseteq E$ be an open subset. Let

$$\begin{array}{rccc} \pi : & E & \longrightarrow & Q \\ & x & \longmapsto & [x]. \end{array}$$

Let $x \in U$, $r > 0$ such that $B(x, r) \subseteq U$. For any $\alpha \in Q$, if

$$\|\alpha - [x]\|_Q = \inf_{y\in\alpha} \|y - x\|_E < r,$$

then, exists $y \in \alpha$ such that $\|y - x\|_E < r$.
Therefore,

$$B([x], r) \subseteq \pi(B(x, r)) \subseteq \pi(U).$$

This means that $\pi(U)$ is open. So $f(U) = g(\pi(U))$ is open. □

Definition 9.9.6 Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector space over a complete non-trivially valued field $(K, |\cdot|)$, $U \subseteq E$ open, $f : U \longrightarrow F$. If $\forall p \in U$, f is n-times differentiable at p, and $\mathrm{D}^n f : U \longrightarrow \mathscr{L}^{(n)}(E, \ldots, E, F)$ is continuous, we say that f is of class $\mathcal{C}^n$.

If $\forall n \in \mathbb{N}$, f is n-times differentiable on U, we say that f is smooth, or of class $\mathcal{C}^\infty$. ($\forall n \in \mathbb{N}$, f is of class $\mathcal{C}^n$.)

Proposition 9.9.7 Let $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ and $(G, \|\cdot\|_G)$ be normed vector space over a complete non-trivially valued field $(K, |\cdot|)$. $U \subseteq E$, $V \subseteq F$ be open subsets, $f : U \longrightarrow V$, $g : V \longrightarrow G$ be mappings. $n \in \mathbb{N}$.
(1) Let $p \in U$. If f is n-times differentiable at p and g is n-times differentiable at $f(p)$, then $g \circ f$ is n-times differentiable at p.
(2) If f is of class $\mathcal{C}^n$ on U and g is of class $\mathcal{C}^n$ on V, then $g \circ f$ is of class $\mathcal{C}^n$ on U.

Proof (induction on n)
$n = 0$, continuity composition.
$n = 1$, differentiability of composition.
$n \geq 2$,

$$\mathrm{D}(g \circ f)(p)(\cdot) = \mathrm{D}g(f(p))(\mathrm{D}f(p)(\cdot))$$

Let

$$\begin{array}{rrcl} \Phi : & \mathscr{L}(F, G) \times \mathscr{L}(E, F) & \longrightarrow & \mathscr{L}(E, G) \\ & (\alpha, \beta) & \longmapsto & \alpha \circ \beta. \end{array}$$

This is a bounded bilinear mapping. $\|\alpha \circ \beta\| \leq \|\alpha\| \cdot \|\beta\|$.

$$(\|\alpha \circ \beta(h)\|_G = \|\alpha(\beta(h))\|_G \leq \|\alpha\| \cdot \|\beta(h)\|_F \leq \|\alpha\|\|\beta\|\|h\|_E)$$

Φ is of class $\mathcal{C}^\infty$.

$$\mathrm{D}(g \circ f) = \Phi(\mathrm{D}g \circ f, \mathrm{D}f).$$

(2) Since $\mathrm{D}g$ and $\mathrm{D}f$ is of class $\mathcal{C}^{n-1}$, we obtain that $\mathrm{D}(g \circ f)$ is of class $\mathcal{C}^{n-1}$, so $g \circ f$ is of class $\mathcal{C}^n$.
(1) If g is n-times differentiable at $f(p)$, $\mathrm{D}g$ is $(n-1)$-times differentiable at $f(p)$.

So $\mathrm{D}g \circ f$ is $(n-1)$-times differentiable at p. $\mathrm{D}f$ is $(n-1)$-times differentiable at p. So $\mathrm{D}(g \circ f)$ is $(n-1)$-times differentiable at p. $\square$

Theorem 9.9.8 Let $(E, \|\cdot\|)$ be a Banach space over a complete non-trivially valued filed $(K, |\cdot|)$. Let

$$\mathrm{GL}(E) := \{\varphi \in \mathscr{L}(E, E) \mid \varphi \text{ is invertible}\}.$$

This set forms a group under $\circ$.
(1) $\forall \varphi \in \mathscr{L}(E, E)$, if $\|\varphi\| < 1$, then $\mathrm{Id}_E + \varphi \in \mathrm{GL}(E)$.
(2) $\mathrm{GL}(E) \subseteq \mathscr{L}(E, E)$ is open.
(3)

$$\begin{array}{rccc} \iota: & \mathrm{GL}(E) & \longrightarrow & \mathrm{GL}(E) \\ & \varphi & \longmapsto & \varphi^{-1} \end{array}$$

is of class $\mathcal{C}^\infty$.

Proof
(1) The series $\sum_{n\in\mathbb{N}}(-1)^n\varphi^n$ converges absolutely since $\|\varphi^n\| \leq \|\varphi\|^n$.
Let η be the limit of $\sum_{n\in\mathbb{N}}(-1)^n\varphi^n$.

$$(\mathrm{Id} + \varphi) \circ \sum_{k=0}^{n} (-1)^k \varphi^k = \mathrm{Id} + (-1)^n \varphi^{n+1}.$$

Taking the limit when $n \to +\infty$, we get $(\mathrm{Id}_E + \varphi) \circ \eta = \mathrm{Id}_\mathrm{E}$. For the same reason, $\eta \circ (\mathrm{Id}_E + \varphi) = \mathrm{Id}_E$.
(2) If $f \in \mathrm{GL}(E)$, $\forall \varphi \in \mathscr{L}(E, E)$ such that

$$\|\varphi\| < \frac{1}{\|f^{-1}\|},\ f + \varphi = f \circ \left(\mathrm{Id}_E + f^{-1} \circ \varphi\right),\ \|f^{-1} \circ \varphi\| \leq \|f^{-1}\| \cdot \|\varphi\| < 1.$$

So $\mathrm{Id}_E + f^{-1} \circ \varphi \in \mathrm{GL}(E)$. Hence $f + \varphi \in \mathrm{GL}(E)$.
(3) Let $f \in \mathrm{GL}(E)$, $\varphi \in \mathscr{L}(E, E)$. $\|\varphi\| \leq \frac{1}{\|f^{-1}\|}$.

$$\begin{aligned}
\iota(f + \varphi) - \iota(f) &= (f + \varphi)^{-1} - f^{-1} \\
&= \left(f \circ \left(\mathrm{Id}_E + f^{-1} \circ \varphi\right)\right)^{-1} - f^{-1} \\
&= \left(\mathrm{Id}_E + f^{-1} \circ \varphi\right)^{-1} \circ f^{-1} - f^{-1} \\
&= \sum_{n\in\mathbb{N}} (-1)^n \left(f^{-1} \circ \varphi\right)^n \circ f^{-1} - f^{-1} \\
&= - f^{-1} \circ \varphi \circ f^{-1} + o(\|\varphi\|)
\end{aligned}$$

since

$$\begin{aligned}&\sum_{n\geq 2}\|(-1)^n\left(f^{-1}\circ\varphi\right)^n\circ f^{-1}\|\\ \leq&\sum_{n\geq 2}\|f^{-1}\|\cdot\left(\|f^{-1}\|\cdot\|\varphi\|\right)^n\\ =&\|\varphi\|^2\left(\|f\|^3\cdot\sum_{n\geq 2}\left(\|f^{-1}\|\cdot\|\varphi\|\right)^{n-2}\right)\\ =&o(\|\varphi\|).\end{aligned}$$

Let

$$\begin{array}{rccc}\Phi: & \mathscr{L}(E,E)^3 & \longrightarrow & \mathscr{L}(E,E)\\ & (\alpha,\beta,\gamma) & \longmapsto & \alpha\circ\beta\circ\gamma.\end{array}$$

bounded 3-linear mapping.

$$\mathrm{D}\iota(f)(\cdot) = -\Phi\left(\iota(f),\cdot,\iota(f)\right).$$

By induction, we obtain that ι is of class $\mathcal{C}^n$ for any $n\in\mathbb{N}$. □

Definition 9.9.9 Let (X,d) be a metric space, $f: X\longrightarrow X$ be a mapping. If exists $\alpha\in\,]0,1[$, such that f is α-Lipschitzian, we say that f is a **contraction**.

Definition 9.9.10 Let $f: X\longrightarrow X$ be a mapping. If $x\in X$ is such that $f(x)=x$, we say that x is a **fixed point** of f.

Theorem 9.9.11 (Banach fixed point theorem) Let (X,d) be a non-empty complete metric space and $f: X\longrightarrow X$ be a contraction. Then f admits a unique fixed point.

Proof
" Uniqueness": Let $\alpha\in\,]0,1[$, such that f is α-Lipschitzian. If a and b are fixed point of f, then $\mathrm{d}(a,b)=\mathrm{d}(f(a),f(b))\leq\alpha\mathrm{d}(a,b)$. So $\mathrm{d}(a,b)=0$, $a=b$.
" Existence": Let $x_0\in X$. For any $n\in\mathbb{N}$, let $x_n=f^n(x_0)$. Then

$$\mathrm{d}(x_n,x_{n+1})=\mathrm{d}(f(x_{n-1}),f(x_n))\leq\alpha\mathrm{d}(x_{n-1},x_n)\leq\cdots\leq\alpha^n\mathrm{d}(x_0,x_1).$$

So

$$\sum_{n\in\mathbb{N}}\mathrm{d}(x_n,x_{n+1})\leq\sum_{n\in\mathbb{N}}\alpha^n\mathrm{d}(x_0,x_1)=\frac{1}{1-\alpha}\mathrm{d}(x_0,x_1)<+\infty.$$

Hence $(x_n)_{n\in\mathbb{N}}$ converges to some $a \in X$.

$$\mathrm{d}(a, f(a)) = \lim_{n\to+\infty} \mathrm{d}(x_n, f(x_n)) = \lim_{n\to+\infty} \mathrm{d}(x_n, x_{n+1}) = 0.$$

So $a = f(a)$. □

Definition 9.9.12 Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over a complete value filed $(K, |\cdot|)$, $U \subseteq E$, $V \subseteq F$ be open subsets, $f : U \longrightarrow V$ be a bijection, $n \in \mathbb{N} \cup \{\infty\}$. If f and f^{-1} are both of class $\mathcal{C}^n$, we say that f is a $\mathcal{C}^n$-diffeomorphism.

Theorem 9.9.13 Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be Banach spaces over $\mathbb{R}$, $U \subseteq E$ open and $f : U \longrightarrow F$ be a mapping of class $\mathcal{C}^n$ ($n \in \mathbb{N} \cup \{\infty\}$). Let $p \in U$. Suppose that $\mathrm{D}f(p) \in \mathscr{L}(E, F)$ is invertible. Then there exists a open neighborhood V of p contained in U, such that $f|_V : V \longrightarrow f(V)$ is a $\mathcal{C}^n$-homeomorphism. Moreover

$$\mathrm{D}f^{-1}(y) = \mathrm{D}f(f^{-1}(y))^{-1}.$$

Proof By replacing f by

$$\tilde{f} : x \longmapsto \mathrm{D}f(p)^{-1}\left(f(p+x) - f(p)\right).$$

We may assume that $E = F$, $p = f(p) = 0$, $\mathrm{D}f(p) = \mathrm{Id}_E$.

$$\mathrm{D}\tilde{f}(0)(h) = \mathrm{D}f(p)^{-1}(\mathrm{D}f(p)(h)) = h, \mathrm{D}\tilde{f}(0) = \mathrm{Id}_E.$$

Let $\mu : U \longrightarrow E$, $\mu(x) = f(x) - x$, $\mathrm{D}\mu(0) = 0$. Since $\mathrm{D}f$ is continuous, so is $\mathrm{D}\mu$.

$$\exists r > 0,\ \forall x \in \overline{B}(0_E, r),\ \|\mathrm{D}\mu(x)\| \leq \frac{1}{2}.$$

So μ is $\frac{1}{2}$-Lipschitzian on $\overline{B}(0_E, r)$ (mean value inequality).

$$\forall (x, y) \in \overline{B}(0_E, r)^2,\ \|f(x) - f(y)\| \geq \|x - y\| \|\mu(x) - \mu(y)\| \geq \frac{1}{2}\|x - y\|.$$

So f is injective on $\overline{B}(0_E, r)$. Let $a \in \overline{B}(0_E, \frac{r}{2})$.

$$\forall x \in \overline{B}(0_E, r),\ \|a - \mu(x)\| \leq \|a\| + \|\mu(x)\| \leq \frac{1}{2}r + \frac{1}{2}r = r.$$

Let

$$\begin{array}{rccc} \nu: & \overline{B}(0,r) & \longrightarrow & \overline{B}(0,r) \\ & x & \longmapsto & a-\mu(x) \end{array}$$

ν is a contraction. By Banach's fixed point theorem,

$$\exists! g(a) \in \overline{B}(0,r),\ \nu(g(a)) = a - \mu(g(a)) = a - f(g(a)).$$

That is $f(g(a)) = a$. Let $W = B(0, \frac{r}{2})$, $V = f^{-1}(W) \cap B(0,r)$, $f|_V : V \longrightarrow W$ is a bijection.

$\forall z \in B(0,r)$, $\mathrm{D}f(z) = \mathrm{Id}_E + \mathrm{D}\mu(z) \in \mathrm{GL}(E)$.

$\forall (x, x_0) \in V \times V, y = f(x), y_0 = f(x_0), y - y_0 = \mathrm{D}f(x_0)(x-x_0) + o\left(\|x - x_0\|\right).$

$$\|x - x_0\| = \|y - y_0 - (\mu(x) - \mu(x_0))\| \leq \|y - y_0\| + \frac{1}{2}\|x - x_0\|,$$

$$\frac{1}{2}\|f^{-1}(y) - f^{-1}(y_0)\| = \frac{1}{2}\|x - x_0\| \leq \|y - y_0\|.$$

So,

$$\mathrm{D}f(x_0)(x - x_0) = y - y_0 + o(\|y - y_0\|),$$

$$\begin{aligned} f^{-1}(y) - f^{-1}(y_0) =& x - x_0 = \mathrm{D}f(x_0)^{-1}(y - y_0) + o(\|y - y_0\|) \\ =& \mathrm{D}f(f^{-1}(y_0))^{-1}(y - y_0) + o(\|y - y_0\|) \end{aligned}$$

Thus,

$$\mathrm{D}f^{-1} = \iota \circ \mathrm{D}f \circ f^{-1}.$$

□

Proposition 9.9.14 Let $n \in \mathbb{N}_{\geq 1}$. Let $(K, |\cdot|)$ be a complete valued field, $(E_i, \|\cdot\|_i)$, $i \in \{1, \ldots, n\}$ be normed vector spaces over K, $(F, \|\cdot\|_F)$ be a Banach space over K. Then, $(\mathscr{L}^{(n)}(E_1, \ldots, E_n, F), \|\cdot\|)$ is a Banach space.

Proof Let $(\varphi_i)_{i \in \mathbb{N}}$ be a Cauchy sequence in $\mathscr{L}^{(n)}(E_1, \ldots, E_n, F)$. For $N \in \mathbb{N}$, let

$$\varepsilon_N := \sup_{(i,j) \in \mathbb{N}_{\geq N}} \|\varphi_i - \varphi_j\|, \ \lim_{N \to +\infty} \varepsilon_N = 0.$$

For any $(x_1, \dots, x_n) \in E_1 \times \cdots \times E_n$, and any $(i, j) \in \mathbb{N}^2_{\geq N}$,

$$\|\varphi_i(x_1, \dots, x_n) - \varphi_j(x_1, \dots, x_n)\| \leq \|\varphi_i - \varphi_j\| \cdot \prod_{l=1}^{n} \|x_l\|_l \leq \varepsilon_N \prod_{l=1}^{n} \|x_l\|_l.$$

So $(\varphi_i(x_1, \dots, x_n))_{i\in\mathbb{N}}$ is a Cauchy sequence in F, hence it converges to some element of F, denoted as $\varphi(x_1, \dots, x_n)$.

Note that φ is a point-wise limit of an n-linear mapping, so it is also n-linear.

$$\begin{aligned}\|\varphi(x_1, \dots, x_n)\|_F &= \lim_{i\to+\infty} \|\varphi_i(x_1, \dots, x_n)\|_F \\ &\leq \limsup_{i\to+\infty} \|\varphi_i\| \cdot \|x_1\|_{E_1} \cdots \|x_n\|_{E_n}\end{aligned}$$
$$\leq \left(\sup_{i\in\mathbb{N}} \|\varphi_i\|\right) \cdot \|x_1\|_{E_1} \cdots \|x_n\|_{E_n}$$

So $\varphi \in \mathscr{L}(E_1, \dots, E_n, F)$.

For fixed $N \in \mathbb{N}$, $\forall (x_1, \dots, x_n) \in E_1 \times \dots E_n$,

$$\begin{aligned}&\|\varphi(x_1, \dots, x_n) - \varphi_N(x_1, \dots, x_n)\|_F \\ =& \lim_{n\to+\infty} \|\varphi_n(x_1, \dots, x_n) - \varphi_N(x_1, \dots, x_n)\|_F \\ \leq& \varepsilon_N \|x_1\| \cdot \|x_n\|.\end{aligned}$$

So $0 \leq \|\varphi - \varphi_N\| \leq \varepsilon_N$. By squeeze theorem

$$\lim_{N\to+\infty} \|\varphi - \varphi_N\| = 0.$$

$\square$

9.10 Uniform Convergence

Definition 9.10.1 Let X be a set, $(Y, \mathscr{T})$ be a topological space, $(f_n)_{n\in\mathbb{N}}$ be a sequence of mappings from X to Y. We say that the sequence $(f_n)_{n\in\mathbb{N}}$ converges **point-wise** to a mapping $f : X \longrightarrow Y$ if for every $x \in X$ the sequence $(f_n(x))_{n\in\mathbb{N}}$ converges to $f(x)$.

Suppose that (Y, d) is a metric space. We say that the sequence $(f_n)_{n\in\mathbb{N}}$ converges **uniformly** to a mapping $f : X \longrightarrow Y$ if

$$\lim_{n\to+\infty} \sup_{x\in X} \mathrm{d}(f_n(x), f(x)) = 0.$$

Remark 9.10.2 Let $f, g : X \longrightarrow Y$ be mappings.

$$\mathrm{d}(f, g) = \sup_{x\in X} \mathrm{d}(f(x), g(x))$$

is a metric. Uniform convergence can be seen as convergence of $(f_n)_{n\in\mathbb{N}}$ with respect to this metric.

Theorem 9.10.3 Let $(X, \mathscr{T}_X)$ be a topological space, (Y, d_Y) be a metric space, $(f_n)_{n\in\mathbb{N}}$ be a sequence of mappings from X to Y that converges uniformly to a mapping $f : X \longrightarrow Y$. If $\forall n \in \mathbb{N}$, f_n is continuous at $p \in X$, then f is continuous at p.

Proof We will prove that for any $\varepsilon > 0$, $f^{-1}(B(f(p), \varepsilon))$ is a neighborhood of p.
Let $n \in \mathbb{N}$ such that

$$\sup_{x\in X} \mathrm{d}(f_n(x), f(x)) < \frac{\varepsilon}{3}.$$

We claim that

$$f_n^{-1}(B(f_n(x), \frac{\varepsilon}{3})) \subseteq f^{-1}(B(f(x), \varepsilon)).$$

Let x be a element of X such that

$$\mathrm{d}(f_n(x), f_n(p)) < \frac{\varepsilon}{3}.$$

One has

$$\mathrm{d}(f(x), f(p)) \leq \mathrm{d}(f(x), f_n(x)) + \mathrm{d}(f_n(x), f_n(p)) + \mathrm{d}(f_n(p), f(p)) < \varepsilon.$$

$\square$

Theorem 9.10.4 Let (X, d_X) and (Y, d_Y) be two metric spaces, $(f_n)_{n\in\mathbb{N}}$ be sequence of uniformly continuous mappings from X to Y. Suppose that $(f_n)_{n\in\mathbb{N}}$ converges uniformly to $f : X \longrightarrow Y$. Then f is uniformly continuous.

Proof Let $\varepsilon > 0$. There exists $n \in \mathbb{N}$ such that

$$\sup_{x\in X} \mathrm{d}(f_n(x), f(x)) < \frac{\varepsilon}{3}.$$

f_n is uniformly continuous, so the exists $\delta > 0$

$$\forall (x, y) \in X \times X,\ \mathrm{d}(x, y) < \delta \Rightarrow \mathrm{d}(f_n(x), f_n(y)) < \frac{\varepsilon}{3}.$$

Therefore, for any $(x, y) \in X \times X$ such that $\mathrm{d}(x, y) < \delta$,

$$\mathrm{d}(f(x), f(y)) \leq \mathrm{d}(f(x), f_n(x)) + \mathrm{d}(f_n(x), f_n(y)) + \mathrm{d}(f_n(y), f(y)) < \varepsilon.$$

So f is uniformly continuous. □

Theorem 9.10.5 Let $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ be normed vector spaces over a complete non-trivially valued field $(K, |\cdot|)$. Let $U \subseteq E$ open, $(f_n)_{n\in\mathbb{N}}$ a sequence of differentiable mappings from U to F. Let $f : U \longrightarrow F$, $g : U \longrightarrow \mathscr{L}(E, F)$ be mappings, $p \in U$. Suppose that
(1) The sequence $(\mathrm{D}f_n)_{n\in\mathbb{N}}$ converges uniformly to g.
(2) $(f_n)_{n\in\mathbb{N}}$ converges uniformly to f.
(3) There exists $N \in \mathbb{N}$ and mapping $\delta : U \longrightarrow \mathbb{R}_{\geq 0}$ such that $\lim_{x\to p} \delta(x) = 0$ and for any $n \in \mathbb{N}_{\geq N}$, any $x \in U$,

$$\|f_n(x) - f_n(p) - \mathrm{D}f_n(p)(x - p)\|_F \leq \delta(x)\|x - p\|_E.$$

Then f is differentiable and $\mathrm{D}f = g$.

Proof For any $n \in \mathbb{N}$, define

$$\varepsilon_n := \sup_{x\in U} \|\mathrm{D}f_n(x) - g(x)\|,\ \mathrm{d}_n := \sup_{x\in U} \|f_n(x) - f(x)\|.$$

One has

$$\begin{aligned}\|f(x)-f(p)-g(p)(x-p)\| \leq &\|(f(x)-f_n(x))-(f(p)-f_n(p))\| \\ &+\|f_n(x)-f_n(p)-\mathrm{D}f_n(p)(x-p)\| \\ &+\|\mathrm{D}f_n(p)(x-p)-g(p)(x-p)\| \\ \leq &2\mathrm{d}_n+\delta(x)\|x-p\|_E+\varepsilon_n\|x-p\|_E.\end{aligned}$$

for sufficiently large n.
Therefore,

$$\limsup_{x\to p}\frac{\|f(x)-f(p)-g(p)(x-p)\|_F}{\|x-p\|_E}\leq 2\varepsilon_n.$$

Taking the limit when $n\to\infty$, we obtain

$$\limsup_{x\to p}\frac{\|f(x)-f(p)-g(p)(x-p)\|_F}{\|x-p\|_E}=0.$$

Namely,

$$f(x)-f(p)-g(p)(x-p)=o(\|x-p\|_E),\ x\to p.$$

□

Theorem 9.10.6 Let $(E,\|\cdot\|_E)$ and $(F,\|\cdot\|_F)$ be normed vector spaces over $\mathbb{R}$, $U\subseteq E$ open, $(f_n)_{n\in\mathbb{N}}$ be sequence of differentiable mappings from U to F, $g:U\longrightarrow\mathscr{L}(E,F)$. We suppose that
(1) $(\mathrm{D}f_n)_{n\in\mathbb{N}}$ converges uniformly to g.
(2) $(f_n)_{n\in\mathbb{N}}$ converges point-wise to $f:U\longrightarrow F$.
Then f is differentiable and $\mathrm{D}f=g$.

Proof Let $p\in U$, for any $(n,m)\in\mathbb{N}\times\mathbb{N}$, for any $n\in\mathbb{N}$,

$$c_{m,n}:=\sup_{x\in U}\|\mathrm{D}f_m(x)-\mathrm{D}f_n(x)\|,\ \varepsilon_n:=\sup_{x\in U}\|\mathrm{D}f_n(x)-g(x)\|.$$

For $r>0$, $B(p,r)\subseteq U$ by the mean value inequality,

$$\|(f_n(x)-f_m(x))-(f_n(p)-f_m(p))\|\leq c_{m,n}\|x-p\|,\ x\in B(p,r).$$

Passing to the limit when $m\to+\infty$, we obtain

$$\|(f_n(x)-f(x))-(f_n(p)-f(p))\|\leq\varepsilon_n\|x-p\|.$$

we have

$$\begin{aligned}\|f(x)-f(p)-g(p)(x-p)\| \leq & \|(f(x)-f_n(x))-(f(p)-f_n(p))\| \\ & +\|f_n(x)-f_n(p)-\mathrm{D}f_n(p)(x-p)\| \\ & +\|\mathrm{D}f_n(p)(x-p)-g(p)(x-p)\|.\end{aligned}$$

$$\limsup_{x\to p}\frac{\|f(x)-f(p)-g(p)(x-p)\|_F}{\|x-p\|_E}\leq 3\varepsilon_n.$$

Taking the limit $n\to+\infty$

$$\lim_{x\to p}\frac{\|f(x)-f(p)-g(p)(x-p)\|_F}{\|x-p\|_E}=0.$$

Namely,

$$f(x)-f(p)-g(p)(x-p)=o(\|x-p\|_E),\ x\to p.$$

□

Proposition 9.10.7 Let $(E,\|\cdot\|_E)$, $(F,\|\cdot\|_F)$ be normed vector spaces over $\mathbb{R}$. Assume that $(F,\|\cdot\|_F)$ is a Banach space, $U\subseteq E$ be a path connected open, $(f_n)_{n\in\mathbb{N}}$ be a sequence of differentiable mappings from U to F. Suppose that
(1) $(\mathrm{D}f_n)_{n\in\mathbb{N}}$ converges uniformly to $g:U\longrightarrow\mathscr{L}(E,F)$.
(2) There exists $p\in U$ such that $(f_n(p))_{n\in\mathbb{N}}$ converges.
Then the sequence $(f_n)_{n\in\mathbb{N}}$ converges point-wise on U to a differentiable mapping $f:U\longrightarrow F$ such that $\mathrm{D}f=g$.

Proof We first treat the case where U is convex.
For any $(n,m)\in\mathbb{N}\times\mathbb{N}$, let

$$c_{m,n}:=\sup_{x\in U}\|\mathrm{D}f_m(x)-\mathrm{D}f_n(x)\|.$$

Let $x\in U$. By the mean value inequality,

$$\|(f_n(x)-f_m(x))-(f_n(p)-f_m(p))\|\leq c_{m,n}\|x-p\|,$$

which leads to

$$\|f_n(x)-f_m(x)\|_F\leq\|f_n(p)-f_m(p)\|_F+c_{m,n}\|x-p\|_E.$$

Therefore $(f_n(x))_{n\in\mathbb{N}}$ is a Cauchy sequence in F (Banach space), so $f_n(x)$ converges in F to some $f(x)$. Now it suffices to use the theorem 9.10.6.

We will now treat the general case. Let $x \in U$. There exists $\gamma : [0,1] \longrightarrow U$ continuous such that $\gamma(0) = p, \gamma(1) = x$. Let I be the set of $t \in [0,1]$ such that $f_n(\gamma(s))$ converges for all $s \in [0,t]$. By definition, I is an interval in $[0,1]$ and $0 \in I$. Therefore, it is of the form $[0,c]$ or $[0,c[$.

Let $B(\gamma(c),r) \subseteq U$. Since γ is continuous, $\gamma^{-1}(B(\gamma(c),r))$ is open in $[0,1]$ and $c \in \gamma^{-1}(B(\gamma(c),r))$. Assume by contraction that $I = [0,c[$, then $I \cap \gamma^{-1}(B(\gamma(c),r)) \neq \varnothing$. There exists $q \in \gamma^{-1}(B(\gamma(c),r)) \cap I$ such that $f_n(q)$ converges. So from the "convex U version" f converges point-wise on $B(\gamma(c),r)$. So $f_n(\gamma(c))$ converges. Contradiction. We deduce that $I = [0,c]$.

If $c \neq 1$, then c is an adherent point of $]c,1]$. $\gamma^{-1}(B(\gamma(c),r))$ open, so there exists $r' > 0$ such that $B(c,r') \subseteq \gamma^{-1}(B(\gamma(c),r))$. In particular, $B(c,r') \cap]c,1]$ is an open interval in $[0,1]$ that continuous. So $I \supseteq]0,c+r']$. Contradiction. Therefore $c = 1$. □

Definition 9.10.8 Let U be a set and $(F, \|\cdot\|)$ be a Banach space over complete valued field $(K, |\cdot|)$. $(f_n)_{n\in\mathbb{N}} \in \left(F^U\right)^{\mathbb{N}}$ be a sequence of mappings from U to F. If

$$\sum_{n\in\mathbb{N}} \sup_{p\in U} \|f_n(p)\|_F < +\infty,$$

then we say that $\sum\limits_{n\in\mathbb{N}} f_n$ **converges normally**.

Proposition 9.10.9 If $\sum f_n$ converges normally, then it converges uniformly.

Proof For any $n \in \mathbb{N}$, let $g_n = \sum_{k=0}^n f_k$. We need to check that the sequence $(g_n)_{n\in\mathbb{N}}$ converges uniformly. For any $x \in U$, $\sum_{n\in\mathbb{N}} \|f_n(x)\| < +\infty$. So $\sum_{n\in\mathbb{N}} f_n(x)$ converges absolutely. In particular, $(g_n(x))_{n\in\mathbb{N}}$ converges to some $g(x)$.

$$\begin{aligned}\|g_n(x) - g(x)\|_F &= \lim_{m\to+\infty} \|g_n(x) - g_m(x)\|_F \\ &\leq \lim_{m\to+\infty} \|f_{n+1}(x) + \cdots + f_m(x)\|_F \\ &\leq \limsup_{m\to+\infty} \sum_{k\geq n+1} \|f_k(x)\|_F \\ &\leq \varepsilon_n.\end{aligned}$$

Let

$$\varepsilon_n = \sum_{k\geq n+1} \sup_{p\in U} \|f_k(p)\|_F, \quad \lim_{n\to+\infty} \varepsilon_n = 0.$$

So,

$$\limsup_{n\to+\infty}\left(\sup_{x\in U}\|g_n(x)-g(x)\|_F\right)=0,$$

namely, $(g_n)_{n\in\mathbb{N}}$ converges to g. □

Proposition 9.10.10 Let $(K,|\cdot|)$ be a complete valued field which is non-trivially valued, $(E,\|\cdot\|)$ be a normed vector space and $(F,\|\cdot\|_F)$ be a Banach space over K. $U\subseteq E$ be an open subset, $(f_n)_{n\in\mathbb{N}}$ be a sequence of differentiable mappings $U\longrightarrow F$ and $p\in U$. Assume that

(1) $\sum\limits_{n\in\mathbb{N}} f_n$ converges normally (uniformly suffices).

(2) $\sum\limits_{n\in\mathbb{N}} \mathrm{D}f_n$ converges normally (uniformly suffices).

(3) $\exists N\in\mathbb{N}$ and mappings $(\delta_n:U\longrightarrow\mathbb{R}_{\geq 0})_{n\in\mathbb{N}_{\geq N}}$ such that

1. $\forall n\in\mathbb{N}_{\geq N}$, $\lim_{x\to p}\delta_n(x)=\delta_n(p)=0$.
2. $\sum_{n\in\mathbb{N}}\delta_n$ converges normally (uniformly suffices).
3. $\forall n\in\mathbb{N}_{\geq N}$, $\forall x\in U$,

$$\|f_n(x)-f_n(p)-\mathrm{D}f_n(p)(x-p)\|_F\leq\delta_n(x)\|x-p\|_E.$$

Let f and g be limits of $\sum\limits_{n\in\mathbb{N}} f_n$ and $\sum\limits_{n\in\mathbb{N}}\mathrm{D}f_n$ respectively. Then f is differentiable at p and $\mathrm{D}f=g$.

Proposition 9.10.11 Let $(E,\|\cdot\|_E)$ be a normed vector space and $(F,\|\cdot\|_F)$ be a Banach space over $\mathbb{R}$. Let $U\subseteq E$ open, and $(f_n:U\longrightarrow F)_{n\in\mathbb{N}}$ be a sequence of mappings $U\longrightarrow F$. Suppose that

(1) $\sum_{n\in\mathbb{N}}\mathrm{D}f_n$ converges normally (uniformly suffices) to some $g:U\longrightarrow\mathscr{L}(E,F)$.

(2) $\sum_{n\in\mathbb{N}}f_n$ converges point-wise to some $f:U\longrightarrow F$.

Then f is differentiable on U and $\mathrm{D}f=g$.

Remark 9.10.12 If U is path connected, one can replace (2) by (2'): $\exists p\in U$, $\sum_{n\in\mathbb{N}}f_n(p)$ converges.

9.11 Power Series

We fix a complete non-trivially valued field $(K, |\cdot|)$, and let $(E, \|\cdot\|_E)$ be a Banach space over K.

Definition 9.11.1 Let $(S_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ and $b \in K$. We call power series **centered at** b with coefficients $(s_n)_{n\in\mathbb{N}}$ the sequence of polynomial mappings.

$$\left((z \in K) \longmapsto \sum_{l=0}^{n} (z-b)^l s_l\right)_{n\in\mathbb{N}}$$

denoted as

$$\sum_{l=0}^{n} (z-b)^l s_l.$$

If $S = \sum_{n\in\mathbb{N}} (z-b)^n s_n$, we denote by $R(S)$ the element

$$\left(\limsup_{n\to+\infty} \|s_n\|^{\frac{1}{n}}\right)^{-1} \in [0, +\infty]$$

called the **convergence radius** of S. $(0^+ := +\infty, (+\infty)^{-1} := 0)$

Proposition 9.11.2 Let $S = \sum_{n\in\mathbb{N}} (z-b)^n s_n$.
(1) $\forall a \in K$, if $|a-b| < R(S)$, then $S(a) := \sum_{n\in\mathbb{N}} (a-b)^n s_n$ converges absolutely.
(2) If $r > 0$ such that $(r^n \|s_n\|)_{n\in\mathbb{N}}$ is bounded, then $R(S) \geq r$.
(3) If $a \in K$ is such that $|a-b| > R(S)$, then $\sum_{n\in\mathbb{N}} (a-b)^n s_n$ diverges.

Proof
(1)

$$\|(a-b)^n s\|^{\frac{1}{n}} = (|(a-b)^n| \cdot \|s_n\|)^{\frac{1}{n}} = |a-b| \cdot \|s_n\|^{\frac{1}{n}}.$$

$$\limsup_{n\to+\infty} \|(a-b)^n s_n\|^{\frac{1}{n}} = |a-b| \cdot \limsup_{n\to+\infty} \|s_n\|^{\frac{1}{n}}.$$

If $|a-b| < R(S)$, then $|a-b| \cdot \limsup_{n\to+\infty} \|s_n\|^{\frac{1}{n}} < 1$. By the root test of Cauchy, $\sum_{n\in\mathbb{N}} (a-b)^n s_n$ converges absolutely.
(2)

$$\|s_n\|^{\frac{1}{n}} = \frac{1}{r} \left(r^n \|s_n\|\right)^{\frac{1}{n}}.$$

Since $(r^n \|s_n\|)_{n\in\mathbb{N}}$ is bounded,

$$\limsup_{n\to+\infty} \left(r^n \|s_n\|\right)^{\frac{1}{n}} \leq 1.$$

So $\limsup_{n\to+\infty} \|s_n\|^{\frac{1}{n}} \le \frac{1}{r}$. So $R(S) \ge r$.
(3) If $|a-b| > R(S)$, then

$$\limsup_{n\to+\infty} \|(a-b)^n s_n\|^{\frac{1}{n}} = |a-b| \cdot \limsup_{n\to+\infty} \|s_n\|^{\frac{1}{n}} > 1.$$

So $\sum_{n\in\mathbb{N}}(a-b)^n s_n$ diverges. □

Proposition 9.11.3 Let $S = \sum_{n\in\mathbb{N}} (z-b)^n s_n$ be a power series.
(1) $\forall r \in \mathbb{R}_{\ge 0}$ such that $r < R(S)$. the series S converges normally on $\overline{B}(b,r)$.
(2) $(a \in B(b, R(S))) \longmapsto S(a) := \sum_{n\in\mathbb{N}} (a-b)^n S_n$ is continuous.

Proof
(1) $\forall a \in \overline{B}(b,r)$,

$$\sum_{n\in\mathbb{N}} \|(a-b)^n s_n\| \le \sum_{n\in\mathbb{N}} r^n \|s_n\| < +\infty$$

since $\limsup_{n\to+\infty} r \cdot \|s_n\|^{\frac{1}{n}} < 1$.
(2) $a \longmapsto S(a)$ is continuous on any $B(b,r)$, $r < R(S)$. Since

$$B(b, R(S)) = \bigcup_{r<R(S)} B(b,r),$$

S is continuous on $B(b, R(S))$. □

Definition 9.11.4 Let $S = \sum_{n\in\mathbb{N}} (z-b)^n s_n$. We define the formal derivative of S as

$$\sum_{n\in\mathbb{N}_{\ge 1}} (z-b)^{n-1}(n s_n).$$

Proposition 9.11.5 Let $S = \sum_{n\in\mathbb{N}}(z-b)^n s_n$ be a formal power series. Let $P \in K[T]$. For any $n \in \mathbb{N}$, let $P(n) := P(n 1_K) \in K$. Let

$$S_p := \sum_{n\in\mathbb{N}} (z-b)^n (P(n) s_n).$$

Then $R(S_p) \ge R(S)$.

Proof We assume that $P \neq 0$, $P(T)$ is of the form

$$C_dT^d + C_{d-1}T^{d-1} + \cdots + C_1T + C_0,\ C_d \neq 0.$$

$$|P(n)| = \mathcal{O}(n^d) = o(r^n),\ \text{for any } r > 1.$$

Hence, $\exists N \in \mathbb{N}$ such that $|P(n)| \leq r^n$, $\forall n \in \mathbb{N}_{\geq N}$.

$$\limsup_{n\to+\infty} \|P(n)s_n\|^{\frac{1}{n}} \leq r \cdot \limsup_{n\to+\infty} \|s_n\|^{\frac{1}{n}}.$$

Taking the limit when $r \to 1$, get

$$\limsup_{n\to+\infty} \|P(n)s_n\|^{\frac{1}{n}} \leq \limsup_{n\to+\infty} \|s_n\|^{\frac{1}{n}}.$$

$$R(S_p) \geq R(S).$$

□

Lemma 9.11.6 Let $(z_0, z) \in K^2$, $n \in \mathbb{N}_{\geq 1}$.

$$z^n - z_0^n - nz_0^{n-1}(z - z_0) = (z - z_0)^2 \sum_{j=0}^{n-2} (n - j - 1)z^j z_0^{n-2-j}.$$

Proof

$$\begin{aligned}
z^n - z_0^n =&(z - z_0) \sum_{i=0}^{n-1} z^i z^{n-1-i}. \\
z^n - z_0^n - nz_0^{n-1}(z - z_0) =&(z - z_0) \sum_{i=0}^{n-1} (z^i z_0^{n-1-i} - z_0^{n-1}) \\
=&(z - z_0) \sum_{i=0}^{n-1} z_0^{n-i-1}(z^i - z_0^i) \\
=&(z - z_0)^2 \sum_{i=1}^{n-1} z_0^{n-1-i} \sum_{j=0}^{i-1} z^j z_0^{i-j-1} \\
=&(z - z_0)^2 \sum_{j=0}^{n-2} \sum_{i=j+1}^{n-1} z^j z_0^{n-2-j} \\
=&(z - z_0)^2 \sum_{j=0}^{n-2} (n - j - 1)z^j z_0^{n-2-j}.
\end{aligned}$$

□

Theorem 9.11.7 Let $\sum_{n\in\mathbb{N}}(z-b)^n s_n$ be a power series and R be its convergence radius. For any $z \in B(b,R)$, let $S(z)$ be the limit of the series. Then the mapping $S : B(b,R) \longrightarrow E$ is differentiable, and its derivative is given by the limit of the power series

$$\sum_{n\in\mathbb{N}_{\geq 1}} (z-b)^{n-1}(ns_n).$$

Proof Let $r < R$, $(z, z_0) \in B(b,r)^2$.

$$\begin{aligned}&\|(z-b)^n s_n - (z_0-b)^n s_n - (z-z_0)(z_0-b)^{n-1} n s_n\| \\ =&|z-z_0|^2 \cdot \|\sum_{j=0}^{n-2}(n-1-j)(z-b)^j(z_0-b)^{n-2-j}s_n\|.\end{aligned}$$

$$\begin{aligned}&\|\sum_{j=0}^{n-2}(n-1-j)(z-b)^j(z_0-b)^{n-2-j}s_n\| \\ \leq&\sum_{j=0}^{n-2}(n-1-j)r^{n-2}\|s_n\| = \frac{n(n-1)}{2}r^{n-2}\|s_n\|.\end{aligned}$$

We know that

$$\sum_{n\in\mathbb{N}} \frac{n(n-1)}{2} r^{n-2}\|s_n\| < +\infty.$$

Therefore, the result follows from the proposition 9.10.10. □

Definition 9.11.8 Let $(a_n)_{n\in\mathbb{N}} \in K^{\mathbb{N}}$ and $(s_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. We call **Cauchy product** of the series $\sum_{n\in\mathbb{N}} a_n$ and $\sum_{n\in\mathbb{N}} s_n$ as the series:

$$\sum_{n\in\mathbb{N}} \left(\sum_{k=0}^{n} a_k s_{n-k}\right).$$

Theorem 9.11.9 (Merterns) Let $(a_n)_{n\in\mathbb{N}} \in K^{\mathbb{N}}$ and $(s_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. Suppose that $\sum_{n\in\mathbb{N}} a_n$ and $\sum_{n\in\mathbb{N}} s_n$ converges to $b \in K$ and $t \in E$ respectively.
(1) If at least one of $\sum_{n\in\mathbb{N}} a_n$, $\sum_{n\in\mathbb{N}} s_n$ converges absolutely, then the their Cauchy product converges to bt.
(2) If both $\sum_{n\in\mathbb{N}} a_n$ and $\sum_{n\in\mathbb{N}} s_n$ converge absolutely, then the Cauchy product also converges absolutely.

Proof

(1) Suppose that $\sum_{n\in\mathbb{N}} a_n$ converges absolutely. For any $n \in \mathbb{N}$, let

$$A_n := \sum_{k=0}^{n} a_k,\ S_n := \sum_{k=0}^{n} s_k.$$

For any $N \in \mathbb{N}$, let

$$t_N = \sum_{n=0}^{N}\left(\sum_{k=0}^{n} a_k s_{n-k}\right) = \sum_{\substack{(k,l)\in\mathbb{N}^2\\ k+l\leq N}} a_k s_l = \sum_{k=0}^{N} a_k S_{N-k} = A_n t + \sum_{k=0}^{N} a_k\left(S_{N-k}-t\right).$$

Then,

$$t_N - bt = (A_N - b)\,t + \sum_{k=0}^{N} a_k\left(S_{N-k}-t\right).$$

Let $\alpha := \sum_{n\in\mathbb{N}} |a_n|$ and for any $n \in \mathbb{N}$ let

$$\varepsilon_n := \sup_{m\in\mathbb{N},\ m\leq n} \|S_m - t\|.$$

For any $l \in \{0,\dots,N\}$, one has

$$\begin{aligned}\|\sum_{k=0}^{N} a_k\left(S_{N-k}-t\right)\| &\leq \sum_{k=0}^{N-l} |a_k|\cdot\|S_{N-k}-t\| + \sum_{k=N-l+1}^{N} |a_k|\cdot\|S_{N-k}-t\|\\ &\leq \varepsilon_l\cdot\alpha + \max_{i\in\{0,\dots,l-1\}}\|S_i - t\|\cdot\sum_{k=N-l+1}^{N}|a_k|.\end{aligned}$$

We get

$$\forall l\in\mathbb{N},\ \limsup_{N\to+\infty}\|\sum_{k=0}^{N} a_k\left(S_{N-k}-t\right)\| \leq \varepsilon_l\alpha.$$

Taking the infimum with respect to l, we get

$$\limsup_{N\to+\infty}\|\sum_{k=0}^{N} a_k\left(S_{N-k}-t\right)\| = 0.$$

We deduce therefore that

$$\lim_{N\to+\infty} t_N = bt.$$

(2) Let

$$\alpha = \sum_{n\in\mathbb{N}} |a_n|,\ \beta = \sum_{n\in\mathbb{N}} \|s_n\|.$$

For any $N \in \mathbb{N}$, one has

$$\sum_{n=0}^{N} \|\sum_{k=0}^{n} a_k s_{n-k}\| \leq \sum_{n=0}^{N}\sum_{k=0}^{n} |a_k| \cdot \|s_n\| \leq \sum_{\substack{(k,l)\in\mathbb{N}^2\\ k\leq N,\ l\leq N}} |a_k|\|s_l\| \leq \alpha \cdot \beta.$$

So the Cauchy product of $\sum_{n\in\mathbb{N}} a_n$ and $\sum_{n\in\mathbb{N}} s_n$ converges absolutely. □

Example 9.11.10 Consider

$$\mathrm{e}^z = \exp(z) := \sum_{n\in\mathbb{N}} \frac{z^n}{n!},\ z \in \mathbb{C}.$$

By thw ratio test of D'Alembert, for any $r > 0$, $\sum_{n\in\mathbb{N}} \frac{r^n}{n!} < +\infty$. e^z is well defined.
Let $\alpha \in \mathbb{C}$,

$$\exp'(\alpha z) = \alpha \exp(\alpha z).$$

We define

$$\cos(z) := \frac{\mathrm{e}^{\mathrm{i}z} + \mathrm{e}^{-\mathrm{i}z}}{2},\ \sin(z) := \frac{\mathrm{e}^{\mathrm{i}z} - \mathrm{e}^{-\mathrm{i}z}}{2\mathrm{i}},$$
$$\cosh(z) := \frac{\mathrm{e}^{z} + \mathrm{e}^{-z}}{2},\ \sinh(z) := \frac{\mathrm{e}^{z} - \mathrm{e}^{-z}}{2}.$$

Proposition 9.11.11 Let $(a, b, z) \in \mathbb{C}^3$, then

$$\exp((a + b)z) = \exp(az)\exp(bz).$$

Proof The Cauchy product of $\displaystyle\sum_{n\in\mathbb{N}} \frac{(az)^n}{n!}$ and $\displaystyle\sum_{n\in\mathbb{N}} \frac{(bz)^n}{n!}$ is $\displaystyle\sum_{n\in\mathbb{N}} \frac{(a+b)^n z^n}{n!}$. Use the theorem of Merterns. □

9.12 Directional Differential

Definition 9.12.1 Let $(K, |\cdot|)$ be a complete non-trivially valued field, and $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over K. Let $U \subseteq K$ open,

$f : U \longrightarrow F$ be a mapping, $p \in U$, $h \in E$. If the limit

$$\lim_{t\to 0} \frac{f(p+th) - f(p)}{t}$$

exists, we say that f admits the **directional derivative** at p along h.

Notation 9.12.2

$$\partial_h f(p) = \lim_{t\to 0} \frac{f(p+th) - f(p)}{t}.$$

Definition 9.12.3 Let $(E_1, \|\cdot\|_1), \ldots, (E_n, \|\cdot\|_n)$ be normed vector spaces, $E := E_1 \times \cdots \times E_n$,

$$\|(s_1, \ldots, s_n)\| = \max_{i\in\{1,\ldots,n\}} \|s_i\|.$$

If $f : U \longrightarrow F$. We say that f has the **i-th partial differential** at $p = (p_1, \ldots, p_n) \in U$, if the mapping

$$x_i \longmapsto f(p_1, \ldots, p_{i-1}, x_i, p_{i+1}, \ldots, p_n)$$

is differentiable at p_i. We denote the existing differential at p_i by

$$\mathrm{D}_i f(p) \in \mathscr{L}(E_i, F).$$

In the case when $E_i = K$,

$$\mathrm{D}_i f(p)(1) := \partial_i f(p) \text{ or } \frac{\partial f}{\partial x_i}(p).$$

Note that

$$\partial_i f(p) = \partial_{\underset{i\text{-th}}{(0,\cdots,1,\cdots,0)}} f(p).$$

Remark 9.12.4 Let $(K, |\cdot|)$ be a complete non-trivially valued filed, $(E_i, \|\cdot\|_i)$, $i \in \{1, \cdots, n\}$, $(F, \|\cdot\|_F)$ be normed vector spaces. $E = E_1 \times \cdots \times E_n$, equipped with the norm $\|\cdot\|$ defined as

$$\|(x_1, \cdots, x_n)\| = \max_{i\in\{1,\cdots,n\}} \|x_i\|_i.$$

Let $U \subseteq E$ be an open subset, $p \in U$, $f : U \longrightarrow F$ be a mapping. If f is differentiable at p, then f has the i-th partial differential at p for $i \in \{1, \cdots, n\}$.

In fact,

$$f(p_1, \cdots, p_i + h_i, \cdots, p_n) = f(p) + \mathrm{D}f(p)(0, \cdots, h_i, \cdots, 0) + o(\|h_i\|_i).$$

$$\mathrm{D}_i f(p)(h_i) = \mathrm{D}f(p)(0, \cdots, h_i, \cdots, 0).$$

$$\mathrm{D}f(p)(h) = \sum_{i=1}^{n} \mathrm{D}f(p)(0, \cdots, h_i, \cdots, 0) = \sum_{i=1}^{n} \mathrm{D}_i f(p)(h_i).$$

Proposition 9.12.5 Let $(E_i, \|\cdot\|_i)$, $i \in \{1, \cdots, n\}$, $(F, \|\cdot\|_F)$ be normed vector spaces over $\mathbb{R}$, with $\dim_{\mathbb{R}}(F) < +\infty$. Let $E = E_1 \times \cdots \times E_n$, equipped with the norm $\|\cdot\|$ defined as

$$\|(x_1, \cdots, x_n)\| = \max_{i \in \{1, \cdots, n\}} \|x_i\|_i.$$

Let $U \subseteq E$ be an open subset, $f : U \longrightarrow F$ be a mapping. Suppose that, for any $i \in \{1, \cdots, n\}$, f has i^{th} partial differential on U, and $\mathrm{D}_i f : U \longrightarrow \mathscr{L}(E_i, F)$ is continuous. Then f is differentiable on U, and

$$\forall p \in U,\ \mathrm{D}f(p)(h_1, \cdots, h_n) = \sum_{i=1}^{n} \mathrm{D}_i f(p)(h_i).$$

Proof We first treat the case where $F = \mathbb{R}$. Let $p \in U$, and $r > 0$ such that $B(p, r) \subseteq U$. Let $h = (h_1, \cdots, h_n) \in B(0, r)$.

$$\begin{aligned} f(p+h) - f(p) = \sum_{i=1}^{n} (&f(p_1 + h_1, \cdots, p_i + h_i + \cdots, p_{i+1}, \cdots, p_n) \\ &- f(p_1 + h_1, \cdots, p_i, \cdots, p_{i+1}, \cdots, p_n)). \end{aligned}$$

By the mean value theorem of Lagrange,

$$\exists (t_1(h), \cdots, t_n(h)) \in \,]0, 1[^n$$

such that

$$f(p+h) - f(p) = \sum_{i=1}^{n} \mathrm{D}_i f(p_1 + h_1, \cdots, p_i + t_i(h) h_i, \cdots, p_{i+1}, \cdots, p_n)(h_i).$$

$$\begin{aligned}
&f(p+h)-f(p)-\sum_{i=1}^{n}\mathrm{D}_i f(p)(h)\\
=&\sum_{i=1}^{n}\mathrm{D}_i f(p_1+h_1,\cdots,p_i+t_i(h)h_i,\cdots,p_{i+1},\cdots,p_n)(h_i)\\
-&\sum_{i=1}^{n}\mathrm{D}_i f(p_1,\cdots,p_i,\cdots,p_{i+1},\cdots,p_n)(h_i)\\
=&o(\|h\|).
\end{aligned}$$

$\square$

Chapter 10

Integral Calculus

10.1 Differential 1-form

Definition 10.1.1 Let $(K, |\cdot|)$ be a complete non-trivially valued field. Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over K. Let $U \subseteq E$ be an open subset. We call **1-form** *on* U *with coefficients in* F any mapping

$$\alpha : U \longrightarrow \mathscr{L}(E, F).$$

If there exists $f : U \longrightarrow F$ differentiable such that $\mathrm{D}f = \alpha$, we say that α is an **exact** 1-form. (Sometimes $\mathrm{D}f$ is also written as $\mathrm{d}f$.)

Definition 10.1.2 We call a complete valued filed **extension** of $(K, |\cdot|)$ any complete valued field $(K', |\cdot|')$ such that $K \subseteq K'$ and $|\cdot| = |\cdot|'|_K$. Let $(F, \|\cdot\|)$ be a normed vector space over K. If $\alpha : U \longrightarrow \mathscr{L}(E, K')$ and $s : U \longrightarrow F$ be mappings, we denote by

$$\alpha \otimes s : U \longrightarrow \mathscr{L}(E, F)$$

be the mapping sending $p \in U$ to

$$(h \in E) \longmapsto \alpha(p)(h)s(p).$$

Note that

$$\|\alpha(p)(h)s(p)\|_F \leq |\alpha(p)(h)|_{K'} \cdot \|s(p)\|_F \leq \|\alpha(p)\| \cdot \|s(p)\|_F \cdot \|h\|_E.$$

If $(F, \|\cdot\|_F) = (K', |\cdot|')$, $\alpha \otimes s$ is also written as αs.

Example 10.1.3 $(K, |\cdot|) = (\mathbb{R}, |\cdot|)$, $K' = \mathbb{C}$, $|x + \mathrm{i}y|' := \sqrt{x^2 + y^2}$.

Example 10.1.4 Let $\varphi \in \mathscr{L}(E, F)$,

$$\begin{array}{rrcl} \mathrm{D}\varphi : & E & \longrightarrow & \mathscr{L}(E, F) \\ & p & \longmapsto & \varphi. \end{array}$$

is a constant mapping.
As a 1-form, it is often written as $\mathrm{d}\varphi$.

Example 10.1.5 $E = K^n$, $x_i : K^n \longrightarrow K$, $(p_1, \cdots, p_n) \longmapsto p_i$. $U \subseteq E$ open, $f : U \longrightarrow K$ differentiable.

$$\mathrm{d}f(p) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p)\mathrm{d}x_i.$$

Example 10.1.6 Let $w \in \mathbb{C}$, $f : \mathbb{R} \longrightarrow \mathbb{C}$, $t \longmapsto \exp(wt)$.

$$\mathrm{d}f(t) = f'(t)\mathrm{d}t = w \exp(wt)\mathrm{d}t.$$

Proposition 10.1.7 Let $(K', |\cdot|)$ be a complete valued extension of $(K, |\cdot|)$, and $(F, \|\cdot\|_F)$ be a normed vector space over K'. Let $(E, \|\cdot\|_E)$ be a normed vector space over K, $U \subseteq E$ be ann open subset. Let $f : U \longrightarrow K'$ and $g : U \longrightarrow F$ be two mappings that are differentiable, then

$$\mathrm{d}(fg) = f\mathrm{d}g + \mathrm{d}f \otimes g.$$

Proposition 10.1.8 Let $(K', |\cdot|')$ be a complete valued extension of $(K, |\cdot|)$. $(E, \|\cdot\|_E)$ be a normed vector space over K, $(F, \|\cdot\|_F)$ be a normed vector space over K'. Let $U \subseteq E$ be an open subset, and $V \subseteq K'$ be an open subset. $f : U \longrightarrow V$, $g : V \longrightarrow F$ be differentiable mappings, then

$$\mathrm{d}(g \circ f) = \mathrm{d}f \otimes (g' \circ f).$$

Proof For $p \in U$ and $h \in E$,

$$\begin{aligned} \mathrm{D}(g \circ f)(p)(h) &= \mathrm{D}g(f(p))(\mathrm{D}f(p)(h)) \\ &= \mathrm{D}f(p)(h) \cdot \mathrm{D}g(f(p))(1) \\ &= \mathrm{D}f(p)(h) \cdot g'(f(p)) \end{aligned}$$

□

10.2 Primitive Functions

Proposition 10.2.1 Let $(E, \|\cdot\|)$ and $(F, \|\cdot\|)$ be normed vector spaces over $\mathbb{R}$ and $U \subseteq E$ be a path connected open subset. If $f : U \longrightarrow F$ is a mapping such that $\mathrm{d}f = 0$, then f is a constant mapping.

Proof Let p and q be elements of U. There exists $\gamma : [0, 1] \longrightarrow U$ continuous and differentiable on $]0, 1[$, such that $\gamma(0) = p$, $\gamma(1) = q$.

$$\|f(p) - f(q)\|_F = \|f(\gamma(0)) - f(\gamma(1))\|_F \leq \sup_{t \in]0,1[} \|\mathrm{D}f(\gamma(t))(\gamma'(t))\|_F = 0.$$

So $f(p) = f(q)$. □

Definition 10.2.2 Let $I \subseteq \mathbb{R}$ be an open interval and $\varphi : I \longrightarrow F$ be a mapping. If there exists $\Phi : I \longrightarrow F$ such that $\Phi' = \varphi$, we say that Φ is a primitive function of φ. We denote by

$$\int \varphi(t)\mathrm{d}t$$

an arbitrary primitive function of φ. By the previous proposition,

$$\int \varphi(t)\mathrm{d}t = \Phi(t) + C.$$

where C is a constant mapping.

Example 10.2.3 Let $w \in \mathbb{C}$,

$$\int \exp(wt)\mathrm{d}t = \begin{cases} \frac{\exp(wt)}{w} + C & , w \neq 0 \\ t + C & , w = 0. \end{cases}$$

Proposition 10.2.4 Let $I \subseteq \mathbb{R}$ be an open interval. Let $g : I \longrightarrow \mathbb{R}$ and $\varphi : I \longrightarrow F$ be mappings having $G : I \longrightarrow \mathbb{R}$ and $\Phi : I \longrightarrow F$ as primitive functions. Then

$$\int G(t)\mathrm{d}\Phi(t) + \int \mathrm{d}G(t) \otimes \Phi(t) = G(t)\Phi(t) + C.$$

or equivalently,

$$\int G(t)\mathrm{d}t \otimes \varphi(t) + \int g(t)\mathrm{d}t \otimes \Phi(t) = G(t)\Phi(t) + C.$$

If $F = \mathbb{R}$ or $F = \mathbb{C}$, the formula can be written as

$$\int G(t)\mathrm{d}\Phi(t) + \int \Phi(t)\mathrm{d}G(t) = G(t)\Phi(t) + C.$$

or

$$\int G(t)\varphi(t)\mathrm{d}t + \int \Phi(t)g(t)\mathrm{d}t = G(t)\Phi(t) + C.$$

Example 10.2.5

$$\int t\mathrm{e}^t\mathrm{d}t = \int t\mathrm{d}(\mathrm{e}^t) = t\mathrm{e}^t - \int \mathrm{e}^t\mathrm{d}t = t\mathrm{e}^t - \mathrm{e}^t + C.$$

Proposition 10.2.6 Let $U \subseteq \mathbb{R}$ be an open subset, $V \subseteq \mathbb{R}$ be an open subset, $f : U \longrightarrow V$ and $g : V \longrightarrow F$ differentiable mappings. One has

$$\int \mathrm{d}f(t) \otimes g'(f(t)) = g(f(t)) + C.$$

Example 10.2.7

$$\int \sin(t)\cos(t)\mathrm{d}t = \int \sin(t)\mathrm{d}(\sin(t)) = \frac{1}{2}\sin(t)^2 + C.$$

10.3 Riesz Space

We fix a set Ω. We equipped $\mathbb{R}^\Omega$ with the partial order $\leq$ as follows:

$$\forall (f,g) \in \mathbb{R}^\Omega \times \mathbb{R}^\Omega,\ f \leq g \Leftrightarrow \forall \omega \in \Omega,\ f(\omega) \leq g(\omega).$$

If $(f_1, \cdots, f_n) \in \left(\mathbb{R}^\Omega\right)^n$, $\inf\{f_1, \cdots, f_n\}$ and $\sup\{f_1, \cdots, f_n\}$ exists.

$$\forall \omega \in \Omega,\ \inf\{f_1, \cdots, f_n\}(\omega) = \min\{f_1(\omega), \cdots, f_n(\omega)\}$$

$$\forall \omega \in \Omega,\ \sup\{f_1, \cdots, f_n\}(\omega) = \max\{f_1(\omega), \cdots, f_n(\omega)\}$$

Definition 10.3.1 We call Riesz space on Ω any vector space S of $\mathbb{R}^\Omega$, such that

$$\forall (f, g) \in S \times S,\ \inf\{f, g\} \in S.$$

Remark 10.3.2 $\forall (f, g) \in S \times S$,

$$\sup\{f, g\} = f + g - \inf\{f, g\} \in S.$$

$$|f| = \sup\{f, 0\} - \inf\{f, 0\} \in S.$$

By induction, $\forall n \in \mathbb{N}_{\geq 1}$, $\forall (f_1, \cdots, f_n) \in S^n$,

$$\inf\{f_1, \cdots, f_n\}, \sup\{f_1, \cdots, f_n\} \subseteq S.$$

$$\forall \omega \in \Omega,\ \sup\{f, g\}(\omega) = \max\{f(\omega), g(\omega)\} = f(\omega) + g(\omega) - \min\{f(\omega), g(\omega)\}.$$

Definition 10.3.3 Let S be a Riesz space on Ω. We call **integral operator** on S any $\mathbb{R}$-linear mapping $I : S \longrightarrow \mathbb{R}$ such that
(1) $\forall (f, g) \in S \times S$, if $f \leq g$, then $I(f) \leq I(g)$.
(2) If $(f_n)_{n\in\mathbb{N}}$ is a decreasing sequence in S, that converges point-wise to constant zero mapping 0, one has

$$\lim_{n\to+\infty} I(f_n) = 0.$$

Example 10.3.4 Let $\Omega = \mathbb{R}$, $\forall A \subseteq \mathbb{R}$, let

$$\begin{aligned} \mathbb{1}_A :\ \mathbb{R} &\longrightarrow \{0, 1\} \\ x &\longmapsto \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases} \end{aligned}$$

Let S be the vector space of $\mathbb{R}^\mathbb{R}$ generated by mappings of the form $\mathbb{1}_{]a,b]}$, $(a \leq b)$
Let $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ be a right continuous mapping,

$$\forall t \in \mathbb{R},\ \varphi(t) = \lim_{\varepsilon>0, \varepsilon\to 0} \varphi(t + \varepsilon).$$

which is increasing. Then $I_\varphi : S \longrightarrow \mathbb{R}$,

$$I_\varphi\left(\sum_{i=1}^n \lambda_i \mathbb{1}_{]a_i,b_i]}\right) := \sum_{i=1}^n \lambda_i \left(\varphi(b_i) - \varphi(a_i)\right)$$

is an integral operator.

Proposition 10.3.5 Let Ω be a set and S be a Riesz space on Ω. An $\mathbb{R}$-linear mapping $I : S \longrightarrow \mathbb{R}$ that satisfies $(f \leq g \Rightarrow I(f) \leq I(g))$ is an integral operator if and only if, for any increasing sequence $(f_n)_{n\in\mathbb{N}}$ in S that converges point-wise to some $f \in S$, one has

$$\lim_{n\to+\infty} I(f_n) = I(f).$$

Proof
"$\Rightarrow$": $(f - f_n)_{n\in\mathbb{N}}$ is decreasing and converges to 0. So

$$\lim_{n\to+\infty} I(f - f_n) = 0.$$

So $\lim_{n\to+\infty} I(f_n) = I(f)$.
"$\Leftarrow$": Let $(f_n)_{n\in\mathbb{N}}$ be a decreasing sequence in S that converges point-wise to 0. Then $(-f_n)_{n\in\mathbb{N}}$ is increasing and converges point-wise to 0. So

$$\lim_{n\to+\infty} I(-f_n) = 0.$$

So, $\lim_{n\to+\infty} I(f_n) = 0$. □

Proposition 10.3.6 Let Ω be a set and S be a Riesz space on Ω and $I : S \longrightarrow \mathbb{R}$ be an integral operator. Let $g \in S$ and $(f_n)_{n\in\mathbb{N}}$ be an increasing sequence in S. If

$$\forall\omega \in \Omega,\ g(\omega) \leq \lim_{n\to+\infty} f_n(\omega),$$

then

$$I(g) \leq \lim_{n\to+\infty} I(f_n).$$

Proof $(\inf\{g, f_n\})_{n\in\mathbb{N}}$ is an increasing sequence in S. It converges to g. Hence,

$$I(g) = \lim_{n\to+\infty} I(\inf\{g, f_n\}) \leq \lim_{n\to+\infty} I(f_n).$$

□

Corollary 10.3.7 Let $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ be increasing sequences in S. Suppose that

$$\forall\omega \in \Omega,\ \lim_{n\to+\infty} f_n(\omega) \leq \lim_{n\to+\infty} g_n(\omega).$$

Then,

$$\lim_{n\to+\infty} I(f_n) \leq \lim_{n\to+\infty} I(g_n).$$

Proof $\forall k \in \mathbb{N}, \forall \omega \in \Omega,$

$$f_k(\omega) \leq \lim_{n\to+\infty} f_n(\omega) \leq \lim_{n\to+\infty} g_n(\omega).$$

So $I(f_k) \leq \lim_{n\to+\infty} I(g_n)$. Taking the limit when $k \to +\infty$, we get

$$\lim_{k\to+\infty} I(f_k) \leq \lim_{n\to+\infty} I(g_n).$$

□

Definition 10.3.8 Let $S^{\uparrow}$ be the set of all mappings $f : \Omega \longrightarrow]-\infty, +\infty]$ that can be written as the point-wise limit of an increasing sequence in S.

Remark 10.3.9
(1) If $f \in S^{\uparrow}$, $\lambda > 0$, then $\lambda f \in S^{\uparrow}$.
(2) If $(f, g) \in S^{\uparrow} \times S^{\uparrow}$, then $f + g \in S^{\uparrow}$, $\inf\{f, g\} \in S^{\uparrow}$, $\sup\{f, g\} \in S^{\uparrow}$.
(3) If $I : S \longrightarrow \mathbb{R}$ is an integral operator, then for any $f \in S^{\uparrow}$ that is written as the point-wise limit of two increasing sequences $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ in S, then

$$\lim_{n\to+\infty} I(f_n) = \lim_{n\to+\infty} I(g_n).$$

We denote by $I(f)$ this limit.

Proposition 10.3.10 Let $(f_n)_{n\in\mathbb{N}} \in (S^{\uparrow})^{\mathbb{N}}$ be an increasing sequence, and f be its point-wise limit. Then $f \in S^{\uparrow}$, and $I(f) = \lim_{n\to+\infty} I(f_n)$ for any operator I.

Proof For any $k \in \mathbb{N}$, let $(g_{k,m})_{m\in\mathbb{N}} \in S^{\mathbb{N}}$ be an increasing sequence in S that converges point-wise to f_k. For any $n \in \mathbb{N}$, let

$$h_n = \sup\{g_{0,n}, g_{1,n}, \cdots, g_{n,n}\} \in S.$$

$(h_n)_{n\in\mathbb{N}}$ is an increasing sequence in S.
$\forall n \in \mathbb{N}, \forall k \in \mathbb{N}, k \leq n$, one has

$$f_n \geq f_k \geq g_{k,n}, \;\; f_n \geq h_n.$$

So,

$$f = \lim_{n\to+\infty} f_n \geq \lim_{n\to+\infty} h_n \geq \lim_{n\to+\infty} g_{k,n} = f_k.$$

This leads to

$$f = \lim_{n\to+\infty} h_n,\ f \in S^{\uparrow}.$$

One has

$$I(f) = \lim_{n\to+\infty} I(h_n) \leq \lim_{n\to+\infty} I(f_n).$$

Moreover, $\forall n \in \mathbb{N}$, $f \geq f_n$, so $I(f) \geq I(f_n)$. Thus leads to

$$I(f) \geq \lim_{n\to+\infty} I(f_n).$$

□

Definition 10.3.11 Let Ω be a set and S be a Riesz space on Ω. We denote by $S^{\downarrow}$ the set of all mappings $f : \Omega \longrightarrow [-\infty, +\infty[$ that can be written as the point-wise limit of a decreasing sequence in S.

Remark 10.3.12
(1) $f \in S^{\downarrow} \Leftrightarrow -f \in S^{\uparrow}$.
(2) If $f \in S^{\downarrow}$, $\lambda > 0$, then $\lambda f \in S^{\downarrow}$.
(3) If $(f, g) \in S^{\downarrow} \times S^{\downarrow}$, then $f + g \in S^{\downarrow}$, $-\inf\{f, g\} \in S^{\downarrow}$, $-\sup\{f, g\} \in S^{\downarrow}$.
(4) If $(f_n)_{n\in\mathbb{N}} \in (S^{\downarrow})^{\mathbb{N}}$ is a decreasing sequence, then

$$\lim_{n\to+\infty} f_n \in S^{\downarrow}.$$

(5) If $I : S \longrightarrow \mathbb{R}$ is an integral operator. For any $f \in S^{\downarrow}$, let

$$I(f) := -I(-f).$$

1. If $(f, g) \in S^{\downarrow} \times S^{\downarrow}$ or $(f, g) \in S^{\uparrow} \times S^{\uparrow}$,

$$f \leq g \Rightarrow I(f) \leq I(g), I(f + g) = I(f) + I(g),$$

$$I(\lambda f) = \lambda I(f), \forall \lambda \in \mathbb{R}\backslash\{0\}.$$

2. If $(f_n)_{n\in\mathbb{N}} \in (S^{\downarrow})^{\mathbb{N}}$ is a decreasing sequence, then

$$I(\lim_{n\to+\infty} f_n) = \lim_{n\to+\infty} I(f_n).$$

Proposition 10.3.13 Let Ω be a set, S be a Riesz space on Ω and $I : S \longrightarrow \Omega$ be an integral operator. For any $(f, g) \in (S^{\uparrow} \cup S^{\downarrow})^2$, if $f \leq g$, then $I(f) \leq I(g)$.

Proof It is suffices to treat the case where $(f, g) \in S^{\uparrow} \times S^{\downarrow}$ or $(f, g) \in S^{\downarrow} \times S^{\uparrow}$.

If $(f, g) \in S^{\uparrow} \times S^{\downarrow}$, then $(-f, g) \in S^{\downarrow} \times S^{\downarrow}$, so $g - f \in S^{\downarrow}$. $I(g - f) = I(g) - I(f) \geq 0$. So $I(f) \leq I(g)$.

If $(f, g) \in S^{\downarrow} \times S^{\uparrow}$, then $(-f, g) \in S^{\uparrow} \times S^{\uparrow}$, so $g - f \in S^{\uparrow}$. $I(g - f) = I(g) - I(f) \leq 0$. So $I(f) \leq I(g)$. □

Definition 10.3.14 Let Ω be a set, S be a Riesz space on Ω, and $I : S \longrightarrow \mathbb{R}$ be an integral operator. Let $f : \Omega \longrightarrow \mathbb{R}$ be a mapping. If

$$\sup_{\substack{l \in S \\ l \leq f}} I(l) = \inf_{\substack{\mu \in S \\ \mu \geq f}} I(\mu).$$

We say that f is **Riemann integrable**.

Let

$$\underline{I}(f) := \sup_{\substack{l \in S^{\downarrow} \\ l \leq f}} I(l),$$

$$\overline{I}(f) := \inf_{\substack{\mu \in S^{\uparrow} \\ \mu \geq f}} I(\mu),$$

then,

$$\underline{I}(f) \leq I(f) \leq \overline{I}(f).$$

If $\underline{I}(f) = \overline{I}(f) \in \mathbb{R}$, we say that f is **Daniell integrable**, and we denote by $I(f)$ the real number $\underline{I}(f) = \overline{I}(f)$.

We denote by $\mathcal{L}^1(I)$ the set of all Daniell integrable mappings from Ω to $\mathbb{R}$. We got a mapping

$$I : \mathcal{L}^1(I) \longrightarrow \mathbb{R}.$$

Lemma 10.3.15 Let Ω be a set, S be a Riesz space on Ω, and $I : S \longrightarrow \mathbb{R}$ be an integral operator.

(1) For any mapping $f : \Omega \longrightarrow \mathbb{R}$,

$$\underline{I}(-f) = -\overline{I}(f), \ \overline{I}(-f) = -\underline{I}(f).$$

In particular,

$$f \in \mathcal{L}^1(I) \Leftrightarrow -f \in \mathcal{L}^1(I).$$

And in this case,

$$-I(f) = I(-f).$$

(2) For any $(f, g) \in \mathbb{R}^\Omega \times \mathbb{R}^\Omega$,

$$\underline{I}(f+g) \geq \underline{I}(f) + \underline{I}(g),\ \overline{I}(f+g) \leq \overline{I}(f) + \overline{I}(g).$$

In particular, if $(f, g) \in \mathcal{L}^1(I) \times \mathcal{L}^1(I)$, then $f + g \in \mathcal{L}^1(I)$, and $I(f + g) = I(f) + I(g)$.
(3) For any $f \in \mathbb{R}^\Omega$ and any $\lambda \in \mathbb{R}_{>0}$,

$$\underline{I}(\lambda f) = \lambda \underline{I}(f),\ \overline{I}(\lambda f) = \lambda \overline{I}(f).$$

In particular, if $f \in \mathcal{L}^1(I)$, then $\lambda f \in \mathcal{L}^1(I)$, and $I(\lambda f) = \lambda I(f)$.
(4) If $(f, g) \in \mathbb{R}^\Omega \times \mathbb{R}^\Omega$ such that $f \leq g$, then

$$\underline{I}(f) \leq \underline{I}(g),\ \overline{I}(f) \leq \overline{I}(g).$$

(5) If $(f : \Omega \longrightarrow \mathbb{R}) \in S^\uparrow \cup S^\downarrow$ such that $I(f) \in \mathbb{R}$, then $f \in \mathcal{L}^1(I)$.

Proof
(1) If $\mu \in S^\uparrow$, $\mu \geq f$, then $-\mu \in S^\downarrow$, $-\mu \leq -f$. So

$$-I(\mu) = I(-\mu) \leq \underline{I}(-f).$$

$$I(\mu) \geq -\underline{I}(-f).$$

Taking $\inf\limits_{\substack{\mu \in S^\uparrow \\ \mu \geq f}}$, we get

$$\overline{I}(f) \geq -\underline{I}(-f).$$

$\forall l \in S^\downarrow$, $l \leq f$, one has $-l \in S^\uparrow$, $-l \geq -f$. So

$$I(-l) \geq \overline{I}(-f),\ I(l) \leq -\overline{I}(-f).$$

Taking $\sup\limits_{\substack{l \in S^\downarrow \\ l \leq f}}$, we get

$$\underline{I}(f) \leq -\overline{I}(-f).$$

Replacing f by $-f$, we get

$$\underline{I}(-f) \geq -\overline{I}(f),\ -\overline{I}(-f) \geq \underline{I}(f).$$

So $-\underline{I}(-f) = \overline{I}(f),\ -\overline{I}(-f) = \underline{I}(f)$.
(2) For any $(l_1, l_2) \in S^{\downarrow} \times S^{\downarrow}$, $l_1 \leq f$, $l_2 \leq g$. One has $l_1 + l_2 \leq f + g$, so

$$\sup_{\substack{(l_1,l_2)\in S^{\downarrow}\times S^{\downarrow}\\ l_1\leq f, l_2\leq g}} I(l_1 + l_2) \leq \underline{I}(f+g).$$

$$\overline{I}(f+g) = -\underline{I}(-f-g) \geq -(\underline{I}(-f) + \underline{I}(-g)) = \overline{I}(f) + \overline{I}(g).$$

If $\overline{I}(f) = \underline{I}(f)$, $\overline{I}(g) = \underline{I}(g)$, one has

$$\overline{I}(f) + \overline{I}(g) = \underline{I}(f) + \underline{I}(g) \leq \underline{I}(f+g) \leq \overline{I}(f+g).$$

$$\overline{I}(f+g) \leq \overline{I}(f) + \overline{I}(g) = I(f) + I(g).$$

(3)

$$\underline{I}(\lambda f) = \sup_{\substack{l\in S\\ l\leq \lambda f}} I(l) = \sup_{\substack{l\in S^{\downarrow}\\ l\leq f}} I(\lambda l) = \lambda \underline{I}(f).$$

$$\overline{I}(\lambda f) = -\underline{I}(\lambda(-f)) = -\lambda \underline{I}(-f) = \lambda \overline{I}(f).$$

(5) Let $f \in S^{\uparrow}$. By definition, $\overline{I}(f) = I(f)$. Moreover, there exists an increasing sequence $(f_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}} \subseteq (S^{\uparrow})^{\mathbb{N}}$ such that

$$I(f) = \lim_{n\to\infty} I(f_n) \leq \underline{I}(f).$$

So,

$$\underline{I}(f) = I(f) = \overline{I}(f).$$

□

Theorem 10.3.16 (Beppo Levi) Let $(f_n)_{n\in\mathbb{N}}$ be a monotone sequence in $\mathcal{L}^1(I)$ such that converges point-wise to a mapping $f : \Omega \longrightarrow \mathbb{R}$. If $\lim_{n\to+\infty} I(f_n) \in \mathbb{R}$, then

$$f \in \mathcal{L}^1(I),\ I(f) = \lim_{n\to+\infty} I(f_n).$$

Proof Suppose that $(f_n)_{n\in\mathbb{N}}$ is increasing. By replacing f_n by $f_n - f_0$ and f by $f - f_0$, we may assume $f_0 = 0$.

Let $\varepsilon > 0$. For any $n \in \mathbb{N}_{\geq 1}$, let $\mu_n \in S^{\uparrow}$ such that $f_n - f_{n-1} \leq \mu_n$ and

$$I(f_n - f_{n-1}) \geq I(\mu_n) - \frac{\varepsilon}{2^n}.$$

$$f_n = \sum_{k=1}^{n} (f_k - f_{k-1}) \leq \mu_1 + \cdots + \mu_n,$$

and

$$I(f_n) = \sum_{k=1}^{n} I(f_k - f_{k-1}) \geq \sum_{k=1}^{n} \left(I(\mu_k) - \frac{\varepsilon}{2^n} \right) \geq I(\mu_1) + \cdots + I(\mu_n) - \varepsilon.$$

Let

$$\mu = \lim_{N \to +\infty} \sum_{k=1}^{N} \mu_k \in S^{\uparrow}.$$

One has $I(\mu) = \sum_{n \in \mathbb{N}} I(\mu_n)$, $\mu \geq \lim_{n \to +\infty} f_n = f$. Let $\alpha = \lim_{n \to +\infty} I(f_n)$, one has

$$\alpha \geq I(\mu) - \varepsilon \geq \overline{I}(f) - \varepsilon.$$

For any $n \in \mathbb{N}$, let $l_n \in S^{\downarrow}$ such that $l_n \leq f_n \leq f$ and $I(l_n) \geq I(f_n) - \varepsilon$. Then

$$\alpha - \varepsilon \liminf_{n \to +\infty} I(l_n) \leq \underline{I}(f).$$

Thus,

$$\alpha - \varepsilon \underline{I}(f) \leq \overline{I}(f) \leq \alpha + \varepsilon.$$

Since ε is arbitrary, we have

$$\underline{I}(f) = \overline{I}(f) = \alpha = \lim_{n \to +\infty} I(f_n).$$

$\square$

Theorem 10.3.17 (Daniell) $\mathcal{L}^1(I)$ forms a Riesz space on Ω, and $I : \mathcal{L}^1(I) \longrightarrow \mathbb{R}$ is an integral operator extending $I : S \longrightarrow \mathbb{R}$.

Proof By the property of $\overline{I}$ and $\underline{I}$, $\mathcal{L}^1(I)$ is a vector subspace of $\mathbb{R}^\Omega$ and $I : \mathcal{L}^1(I) \longrightarrow \mathbb{R}$ is an $\mathbb{R}$-linear mapping. Moreover, if $f \leq g$, then $I(f) \leq I(g)$. Let $(f_1, f_2) \in \mathcal{L}^1(I)^2$, $\forall \varepsilon > 0$, $\exists (l_1, l_2) \in \left(S^{\downarrow}\right)^2$, $\exists (\mu_1, \mu_2) \in \left(S^{\uparrow}\right)^2$,

$$l_i \leq f_i \leq \mu_i, \ i \in \{1, 2\} \text{ and } I(\mu_i - l_i) \leq \frac{\varepsilon}{2}.$$

Then,

$$\inf\{l_1, l_2\} \leq \inf\{f_1, f_2\} \leq \inf\{\mu_1, \mu_2\},$$

and

$$\inf\{\mu_1, \mu_2\} - \inf\{l_1, l_2\} \leq (\mu_1 - l_1) + (\mu_2 - l_2).$$

Suppose that $\mu_1(\omega) \leq \mu_2(\omega)$, $l_2(\omega) \leq l_1(\omega)$. LFS$=\mu_1(\omega) - l_2(\omega)$, $\mu_2(\omega) \geq \mu_1(\omega) \geq l_1(\omega)$,

$$I\left(\inf\{\mu_1, \mu_2\} - \inf\{l_1, l_2\}\right) \leq I(\mu_1 - l_1) + I(\mu_2 - l_2) \leq \varepsilon.$$

By Beppo Levi's theorem, if $(f_n)_{n\in\mathbb{N}}$ is an increasing sequence in $\mathcal{L}^1(I)$ that converges to some $f \in \mathcal{L}^1(I)$. One has $I(f) = \lim_{n\to+\infty} I(f_n)$. □

Remark 10.3.18 If $f \in \mathcal{L}^1(I)$, then $|f| \in \mathcal{L}^1(I)$.

Theorem 10.3.19 (Fatou's lemma) Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{L}^1(I)$. Assume that there exists $g \in \mathcal{L}^1(I)$ such that $\forall n \in \mathbb{N}$, $f_n \geq g$. Then

$$\liminf_{n\to+\infty} f_n \in \mathcal{L}^1(I),$$

and

$$I\left(\liminf_{n\to+\infty} f_n\right) \leq \liminf_{n\to+\infty} I(f_n).$$

Moreover, when $\liminf_{n\to+\infty} I(f_n) < +\infty$ and $\liminf_{n\to+\infty} f_n$ takes finite values, then

$$\liminf_{n\to+\infty} f_n \in \mathcal{L}^1(I).$$

Proof For any $n \in \mathbb{N}$, let g_n be

$$\inf_{k\in\mathbb{N}} f_{n+k} = \lim_{k\to+\infty} \inf\{f_n, f_{n+1}, \cdots, f_{n+k}\} \geq g.$$

$$I(f_n) \geq \lim_{k\to+\infty} I\left(\inf\{f_n, \cdot, f_{n+k}\}\right) \geq I(g).$$

By Beppo Levi's theorem, $g_n \in \mathcal{L}^1(I)$, and $I(g_n) \leq I(f_n)$. The sequence $(g_n)_{n\in\mathbb{N}}$ is increasing and converges point-wise to $\liminf_{n\to+\infty} f_n$. So $\liminf_{n\to+\infty} f_n \in \mathcal{L}^1(I)^{\uparrow}$, and

$$I\left(\liminf_{n\to+\infty} f_n\right) = \lim_{n\to+\infty} I(g_n) \leq \liminf_{n\to+\infty} I(f_n).$$

If $\liminf_{n\to+\infty} I(f_n) < +\infty$, then $I\left(\liminf_{n\to+\infty} f_n\right) < +\infty$. By Beppo Levi's theorem,

$$\liminf_{n\to+\infty} f_n \in \mathcal{L}^1(I).$$

□

Theorem 10.3.20 (Dominated convergence theorem, Lebesgue) Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{L}^1(I)$ that convergence pointwise to a mapping $f : \Omega \longrightarrow \mathbb{R}$. Assume that there exists $g \in \mathcal{L}^1(I)$ such that

$$\forall n \in \mathbb{N}, |f_n| \leq g,$$

then,

$$f \in \mathcal{L}^1(I) \text{ and } I(f) = \lim_{n\to+\infty} I(f_n).$$

Proof

$$f_n \geq g_n, \; -f_n \geq -g_n, \forall n \in \mathbb{N}.$$

By Fatou's lemma,

$$I(\lim_{n\to+\infty} f_n) \leq \lim_{n\to+\infty} I(f_n),$$

$$I(\lim_{n\to+\infty} (-f_n)) \leq \liminf_{n\to+\infty} I(-f_n) = -\limsup_{n\to+\infty} I(f_n),$$

$$-I(g) \leq \limsup_{n\to+\infty} I(f_n) \leq I(\lim_{n\to+\infty} f_n) \leq \liminf_{n\to+\infty} I(f_n) \leq I(g).$$

So $(I(f_n))_{n\in\mathbb{N}}$ converges to $I\left(\lim_{n\to+\infty} f_n\right) \in \mathbb{R}$. Hence,

$$\lim_{n\to+\infty} f_n \in \mathcal{L}^1(I).$$

$\square$

10.4 Convexity*

Definition 10.4.1 Let E be a vector space over $\mathbb{R}$, $U \subseteq E$ convex. We say that the mapping $f : U \longrightarrow \mathbb{R}$ is **convex** if the **epigraph**

$$\Gamma_+(f) := \{(x, a) \in U \times \mathbb{R} \mid f(x) \leq a\}$$

is convex in $E \times \mathbb{R}$.
We say that $f : U \longrightarrow \mathbb{R}$ is **concave** if its **hypergraph**

$$\Gamma_-(f) := \{(x, a) \in U \times \mathbb{R} \mid f(x) \geq a\}$$

is convex in $E \times \mathbb{R}$.

Proposition 10.4.2 Let E be a vector space over $\mathbb{R}$, $U \subseteq E$ convex, and $f : U \longrightarrow \mathbb{R}$ a mapping. Then the following conditions are equivalent:
(1) f is convex.
(2) For any $(x, y) \in U \times U$, and $t \in [0, 1]$,

$$f(tx + y(1-t)) \leq tf(x) + y(1-t)f(y).$$

Proof
(1)⇒(2): Note that $((x, f(x)), (y, f(y))) \in \Gamma^2_+(f)$, $(x, y) \in U^2$.

$$t(x, f(x)) + (1-t)(y, f(y)) = (tx + y(1-t), tf(x) + (1-t)f(y)) \in \Gamma_+(f).$$

Hence,
$$f(tx + y(1-t)) \leq tf(x) + (1-t)f(y).$$
(2)⇒(1): Let $((x, a), (y, b)) \in \Gamma^2_+(f)$, then $a \geq f(x)$, $b \geq f(y)$. Let $t \in [0, 1]$, then
$$ta + (1-t)b \geq tf(x) + (1-t)f(y) \geq f(tx + (1-t)y).$$
Hence,
$$(tx + (1-t)y, ta + (1-t)b) \in \Gamma_+(f).$$

□

Proposition 10.4.3 Let E be a vector space over $\mathbb{R}$, $U \subseteq E$ convex, and $f : U \longrightarrow \mathbb{R}$ a mapping. Then the following conditions are equivalent:
(1) f is concave.
(2) For any $(x, y) \in U \times U$, and $t \in [0, 1]$,

$$f(tx + y(1-t)) \geq tf(x) + y(1-t)f(y).$$

Proof
(1)⇒(2): Note that $((x, f(x)), (y, f(y))) \in \Gamma^2_-(f)$, $(x, y) \in U^2$.

$$t(x, f(x)) + (1-t)(y, f(y)) = (tx + y(1-t), tf(x) + (1-t)f(y)) \in \Gamma_-(f).$$

Hence,
$$f(tx + y(1-t)) \geq tf(x) + (1-t)f(y).$$
(2)⇒(1): Let $((x, a), (y, b)) \in \Gamma^2_-(f)$, then $a \leq f(x)$, $b \leq f(y)$. Let $t \in [0, 1]$, then
$$ta + (1-t)b \leq tf(x) + (1-t)f(y) \leq f(tx + (1-t)y).$$

Hence,

$$(tx + (1-t)y, ta + (1-t)b) \in \Gamma_-(f).$$

□

Proposition 10.4.4 Let E be a vector space over $\mathbb{R}$, $U \subseteq E$ convex, and $f : U \longrightarrow \mathbb{R}$ a mapping. $(f_i)_{i\in I}$ is a family of linear forms on U. ($f_i : E \longrightarrow \mathbb{R}$ linear.) $(c_i)_{i\in I}$ is a family of real numbers. If

$$\forall p \in U, f(p) = \sup_{i\in I}(f_i(p) + c_i),$$

then, f is convex.

Proof Let $(x, y) \in U^2$, $t \in [0, 1]$, then for any $i \in I$,

$$f_i(tx+(1-t)y)+c_i = t(f_i(x)+c_i)+(1-t)(f_i(y)+c_i) \leq tf(x)+(1-t)f(y).$$

Taking the supremum with respect to i, we obtain

$$f(tx + y(1-t)) \leq tf(x) + (1-t)f(y).$$

□

Proposition 10.4.5 Let $(E, \|\cdot\|)$ be a normed vector space over $\mathbb{R}$, $U \subseteq E$ be a convex open subset, $f : U \longrightarrow \mathbb{R}$ be a differentiable mapping. Then f is convex if and only if

$$\forall (p, x) \in U^2,\ f(x) \geq f(p) + \mathrm{D}f(p)(x-p).$$

Moreover, when f is convex, then

$$\forall x \in U,\ f(x) = \sup_{p\in U}\left(f(p) + \mathrm{D}f(p)(x-p)\right).$$

Proof For any $p \in U$, we define

$$\begin{array}{rccl} g_p : & U & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(p) + \mathrm{D}f(p)(x-p). \end{array}$$

We have that $f(p) = g_p(p)$.

$$\forall (p, x) \in U^2,\ f(x) \geq g_p(x) \Rightarrow f = \sup_{p\in U} g_p.$$

By proposition 10.4.4, f is convex.

Conversely, assume that f is convex, $(p, x) \in U^2$, $t \in [0, 1]$,

$$f(tx+(1-t)p) = f(p+t(x-p)) \leq tf(x)+(1-t)f(p) = f(p)+t(f(x)-f(p)).$$

f is differentiable at p,

$$f(p + t(x - p)) = f(p) + t\mathrm{D}f(p)(x - p) + o(|t|).$$

Taking the limit when $t \to 0$, we get

$$f(x) - f(p) \geq \mathrm{D}f(p)(x - p).$$

□

Definition 10.4.6 Let E be a vector space over $\mathbb{R}$. **Bilinear form** on E is a bilinear mapping from $E \times E$ to $\mathbb{R}$. Let $\varphi : E \times E \longrightarrow \mathbb{R}$ be a symmetric bilinear form.
If

$$\forall x \in E,\ \varphi(x, x) \geq 0,$$

we say that φ is **semipositive**.
If

$$\forall x \in E \backslash \{0\},\ \varphi(x, x) > 0,$$

we say that φ is **positive define**.

Example 10.4.7 Let $(x_1, \cdots, x_n)$ and $(y_1, \cdot, y_n)$ be elements of $\mathbb{R}^n$,

$$((x_1, \cdots, x_n), (y_1, \cdots, y_n)) \longmapsto \sum_{i=1}^{n} x_i y_i$$

is a linear bilinear positive define form on $\mathbb{R}^n$.

Definition 10.4.8 Let E be a vector space over $\mathbb{R}$, $\varphi : E \times E \longrightarrow \mathbb{R}$ be a symmetric bilinear form.

$$\ker(\varphi) := \{x \in E \mid \forall y \in E,\ \varphi(x, y) = 0\}$$

is the intersection of $\ker(\varphi(\cdot, y))$ over all $y \in E$.

The **isotropic cone** of φ is the set of $x \in E$ such that $\varphi(x, x) = 0$. $\ker(\varphi)$ is contained in the isotropic cone of φ.

Proposition 10.4.9 Let E be a vector space over $\mathbb{R}$, $\varphi : E \times E \longrightarrow \mathbb{R}$ be a symmetric bilinear form. If φ is semipositive, then $\ker(\varphi)$ is equal to the isotropic cone of φ.

Proof It is suffices to show that any element y of the isotropic cone of φ is in $\ker(\varphi)$.

Let $x \in E$, $t \in \mathbb{R}$,

$$\varphi(x+ty, x+ty) = \varphi(x,x) + 2t\varphi(x,y) + t^2\varphi(y,y) \geq 0.$$

Since $\varphi(y,y) = 0$, we obtain

$$\forall t \in \mathbb{R},\ \varphi(x,x) + 2t\varphi(x,y) \geq 0,$$

$$\forall -t \in \mathbb{R},\ \varphi(x,x) - 2t\varphi(x,y) \geq 0.$$

Thus, for any $t \in \mathbb{R}$,

$$(\varphi(x,x) + 2t\varphi(x,y))\,(\varphi(x,x) - 2t\varphi(x,y)) = \varphi(x,x)^2 - 4t^2\varphi(x,y)^2 \geq 0.$$

Take the limit $|t| \to +\infty$, we obtain, $\varphi(x,y) = 0$. $\square$

Theorem 10.4.10 (Cauchy-Schwartz) Let E be a vector space over $\mathbb{R}$, $\varphi : E \times E \longrightarrow \mathbb{R}$ be a semipositive, bilinear form. For any $(x,y) \in E \times E$,

$$\varphi(x,y)^2 \leq \varphi(x,x)\varphi(y,y).$$

The equality holds if and only if $\varphi(y-x, h) = 0$ for any $h \in E$.

Proof First, we show that if $[x] = h[y]$ in $E/\ker(\varphi)$ then $\varphi(x,y)^2 = \varphi(x,x)\varphi(y,y)$.
We have

$$\{x - ah, y - bh\} \subseteq \ker\varphi.$$

$$\varphi(x,y) = \varphi((x-ah) + ah, (y-bh) + bh) = \varphi(ah, bh) = ab\varphi(h,h).$$

$$\varphi(x,x) = a^2\varphi(h,h),\ \varphi(y,y) = b^2\varphi(h,h).$$

Hence,

$$\varphi(x,y)^2\varphi(x,x)\varphi(y,y).$$

We know if $\varphi(y,y) = 0$, then $y \in \ker\varphi$. In this case, $[y] = 0$. So $[x], [y]$ are colinear in $E/\ker\varphi$.

Assume that $\varphi(y, y) \neq 0$, $t \in \mathbb{R}$,

$$\varphi(x + ty, x + ty) = t^2\varphi(y, y) + \varphi(x, x) + 2t\varphi(x, y) \geq 0.$$

Take $t = -\frac{\varphi(x,y)}{\varphi(y,y)}$, we obtain

$$\varphi(x, y)^2 \leq \varphi(x, x)\varphi(y, y).$$

If the equality holds, then $\varphi(x + ty, x + ty) = 0$, for $t = -\frac{\varphi(x,y)}{\varphi(y,y)}$ and hence $x + ty \in \ker \varphi$. □

Theorem 10.4.11 Let $(E, \|\cdot\|)$ be a normed vector space over $\mathbb{R}$, $U \subseteq E$ be an open convex subset, $f : U \longrightarrow \mathbb{R}$ be a second-order differentiable mapping. If $\mathrm{D}^2 f(p)$ is semipositive for any p, then f is convex.

Proof Let $(p, x) \in U^2$, we define

$$\begin{array}{rccl} g: & [0,1] & \longrightarrow & \mathbb{R} \\ & t & \longmapsto & f(tx + (1-t)p). \end{array}$$

Then,

$$g'(t) = \mathrm{D}f(p + t(x - p))(x - p),\ g''(t) = \mathrm{D}^2 f(p + t(x - p))(x - p, x - p) \geq 0.$$

By Taylor-Lagrange, there exists $\xi \in [0, 1]$,

$$g(1) - g(0) = g'(0) + \xi g''(\xi) \leq g'(0) = \mathrm{D}f(p)(x - p).$$

So $f(x) - f(p) \geq \mathrm{D}f(p)(x - p)$. So f is convex. □

10.5 Semirings

Definition 10.5.1 Let Ω be a set. We call semiring on Ω any $\mathcal{C} \subseteq \mathscr{P}(\Omega)$ that satisfies
(1) $\varnothing \in \mathcal{C}$.
(2) $\forall (A, B) \in \mathcal{C} \times \mathcal{C}$, $A \cap B \in \mathcal{C}$.
(3) $\forall (A, B) \in \mathcal{C} \times \mathcal{C}$, there exists a family $C_1, \cdots, C_n$ of pairwise disjoint sets in

$\mathcal{C}$ such that
$$B\backslash A = \bigcup_{i=1}^{n} C_i.$$

Example 10.5.2 $\Omega = \mathbb{R}$, $\mathcal{C} = \{]a,b] \mid (a,b) \in \mathbb{R}^2, a \leq b\}$.
(1) $\varnothing =]0,0] \in \mathcal{C}$.
(2) $]a,b] \cap]c,d] \neq \varnothing \Leftrightarrow c < b$ $(a \leq c)$. When $c < b$, $]a,b] \cap]c,d] =]c,b]$.
(3) $B\backslash A = B\backslash(A \cap B)$. We may assume $A \subseteq B$. If $A =]a,b]$, $B =]c,d]$, $A \subseteq B$ implies $c \leq a$, $b \leq d$.

Proposition 10.5.3 Let Ω be a set and $\mathcal{C}$ be a semiring on Ω.
(1) Let $B \in \mathcal{C}$. Let $A_1, \cdots, A_n$ be sets in $\mathcal{C}$. Then $B\backslash(A_1 \cup \cdots \cup A_n)$ can be written as the union of a finite family of pairwise disjoint sets in $\mathcal{C}$.
(2) Let Θ be a finite subset of $\mathcal{C}$. There exists a finite family Φ of pairwise disjoint sets in $\mathcal{C}$ such that each element of Θ can be written as the union of some elements of Φ.
(3) Let $\mathcal{A}$ be the set
$$\{A \in \mathscr{P}(\Omega) \mid \exists n \in \mathbb{N}, \exists (A_1, \cdots, A_n) \in \mathcal{C}^n, A = A_1 \cup \cdots \cup A_n\}.$$
Then any $A \in \mathcal{A}$ can be written as the union of a finite family of pairwise disjoint sets in $\mathcal{C}$. In particular, $\forall (A, A') \in \mathcal{A}^2$, $\{A \cup A', A \cap A', A\backslash A'\} \subseteq \mathcal{A}$.

Proof
(1) We reason by induction on n. The case where $n = 0$ is trivial. Suppose that $B\backslash(A_1, \cup \cdots, \cup A_{n-1}) = \bigcup_{i=1}^{n} C_i$, where $C_1, \cdots, C_m$ are pairwise disjoint sets in $\mathcal{C}$. Then
$$B\backslash(A_1, \cup \cdots, \cup A_n) = \bigcup_{i=1}^{n} C_i \backslash A_n.$$
Each $C_i\backslash A_n$ is of the form $\bigcup_{j=1}^{d_i} D_{ij}$ with $D_{i,1}, \cdots, D_{i,d_i}$ in $\mathcal{C}$, pairwise disjoint. So
$$B\backslash(A_1, \cup \cdots, \cup A_n) = \bigcup_{i=1}^{n} \bigcup_{j=1}^{d_i} D_{ij}.$$

(2) Suppose that $\Theta = \{B_1, \cdots, B_n\}$. For any $i \in \{1, \cdots, n\}$, one has

$$B_i = \bigcup_{i \in J \subseteq \{1,\cdots,n\}} \left(\bigcap_{j \in J} B_j \right) \setminus \left(\bigcup_{k \in \{1,\cdots,n\} \setminus J} B_k \right).$$

For any $J \subseteq \{1, \cdots, n\}$, $J \neq \varnothing$, we let

$$B_j := \left(\bigcap_{j \in J} B_j \right) \setminus \left(\bigcup_{k \in \{1,\cdots,n\} \setminus J} B_k \right) = \left(\bigcap_{j \in J} B_j \right) \bigcap \left(\bigcap_{k \in \{1,\cdots,n\} \setminus J} \complement_\Omega B_k \right).$$

$(B_J)_{\substack{J \subseteq \{1,\cdots,n\} \\ J \neq \varnothing}}$ are pairwise disjoint. By (1), each B_J is the union of fa finite family of pairwise disjoint elements $C_{J,1}, \cdots, C_{J,d_J}$ in $\mathcal{C}$. Let

$$\Phi = \{C_{J,l} \mid l \in \{1, \cdots, d_J\}, J \subseteq \{1, \cdots, n\}, J \neq \varnothing\}.$$

(3) By (2), there exists a finite subset Φ of pairwise disjoint elements of $\mathcal{C}$ such that each A_i is the union of some sets in Φ. Then

$$A = \bigcup_{\substack{C \in \Phi \\ C \subseteq A}} C.$$

□

Proposition 10.5.4 Let Ω be a set and $\mathcal{C}$ be a semiring on Ω. Let S be the vector subspace of $\mathbb{R}^\Omega$ generated by mappings of the form $\mathbb{1}_A$, $A \in \mathcal{C}$.
(1) Any pair $(f, g) \in S^2$ can be written as

$$f = \sum_{i=1}^{n} a_i \mathbb{1}_{C_i}, \ g = \sum_{i=1}^{n} b_i \mathbb{1}_{C_i},$$

where $n \in \mathbb{N}$, $(a_1, \cdots, a_n) \in \mathbb{R}^{\mathbb{N}}$, $(b_1, \cdots, b_n) \in \mathbb{R}^{\mathbb{N}}$, $(C_1, \cdots, C_n) \in \mathcal{C}^n$, pairwise disjoint.
(2) S is a Riesz space.

Proof
(1) By definition, f and g are of the form

$$f = \sum_{A \in \Theta_f} \lambda_A \mathbb{1}_A, \ g = \sum_{B \in \Theta_g} \mu_B \mathbb{1}_B,$$

where Θ_f and Θ_g are finite subsets of $\mathcal{C}$, λ_A and μ_B are real numbers. Let $\Theta = \Theta_f \cup \Theta_g$. There is a subset $\Phi \subseteq \mathcal{C}$ consisting of pairwise disjoint sets, such that element of Θ can be written as the union of some sets in Φ.

Suppose that $\Phi = \{C_1, \cdots, C_n\}$. Then

$$f = \sum_{i=1}^{n} \left(\sum_{\substack{A \in \Theta_f \\ A \cap C_i \neq \varnothing}} \lambda_A \right) \mathbb{1}_{C_i}, \; g = \sum_{i=1}^{n} \left(\sum_{\substack{B \in \Theta_g \\ B \cap C_i \neq \varnothing}} \mu_B \right) \mathbb{1}_{C_i}.$$

(2) If

$$f = \sum_{i=1}^{n} a_i \mathbb{1}_{C_i}, \; g = \sum_{i=1}^{n} b_i \mathbb{1}_{C_i},$$

with $C_1, \cdots, C_n$ in $\mathcal{C}$ pairwise disjoint, then

$$\inf\{f, g\} = \sum_{i=1}^{n} \min\{a_i, b_i\} \mathbb{1}_{C_i} \in S.$$

□

10.6 σ-additive Functions

Definition 10.6.1 Let Ω be a set, $\mathcal{C} \subseteq \mathscr{P}(\Omega)$, $\mu : \mathcal{C} \longrightarrow \mathbb{R}_{\geq 0}$ be a mapping. We say that μ is **additive** if for any finite family $(A_n)_{i=1}^{n}$ of pairwise disjoint sets in $\mathcal{C}$ such that $A_1 \cup \cdots \cup A_n \in \mathcal{C}$. One has

$$\mu(A_1 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \mu(A_i).$$

Remark 10.6.2 If $\varnothing \in \mathcal{C}$, then $\varnothing = \varnothing \cup \varnothing$. So $\mu(\varnothing) = 2\mu(\varnothing)$, that means $\mu(\varnothing) = 0$.

Example 10.6.3 Let $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ right continuous and increasing.

$$\mathcal{C} = \{]a, b] \mid (a, b) \in \mathbb{R}^2, a \leq b\}.$$

We define

$$\mu_\varphi : \mathcal{C} \longrightarrow \mathbb{R}_{\geq 0}, \; \mu_\varphi(]a, b]) = \varphi(b) - \varphi(a).$$

If $a_0 \leq \cdots \leq a_n$ are real numbers.

$$\mu_\varphi\left(]a_0, a_n]\right) = \varphi(a_n) - \varphi(a_0) = \sum_{i=1}^{n} \varphi(a_i) - \varphi(a_{i-1}) = \sum_{i=1}^{n} \mu_\varphi\left(]a_{i-1}, a_i]\right).$$

Therefore, μ_φ is additive.

Proposition 10.6.4 Let Ω be a set, $\mathcal{C}$ be a semiring on Ω, $\mu : \mathcal{C} \longrightarrow \mathbb{R}_{\geq 0}$ be a additive mapping, S be the vector subspace of $\mathbb{R}^\Omega$ generated by $\mathbb{1}_A$, where $A \in \mathcal{C}$.
(1) There exists a unique $\mathbb{R}$-linear mapping $I : S \longrightarrow \mathbb{R}$ such that $I(\mathbb{1}_A) = \mu(A)$ for any $A \in \mathcal{C}$.
(2) Let

$$\mathcal{A} = \{A \in \mathscr{P}(\Omega) \mid \exists n \in \mathbb{N}, \exists (A_1, \cdots, A_n) \in \mathcal{C}^n, A = A_1 \cup \cdots \cup A_n\}.$$

Then μ extends in a unique way to an additive mapping from $\mathcal{A}$ to $\mathbb{R}_{\geq 0}$.

Proof
(1) If $I : S \longrightarrow \mathbb{R}$ exists, then it is unique since S is generated by $\mathbb{1}_A$, $A \in \mathcal{C}$. ($\forall f \in S$, f is of form $\sum_{i=1}^n a_i \mathbb{1}_{A_i}$, $A_i \in \mathcal{C}$, $a_i \in \mathbb{R}$. $I(f)$ should be $\sum_{i=1}^n a_i \mu(A_i)$.) It remains to check that such I is well defined. Suppose that $f \in S$ can be written as

$$f = \sum_{A \in \Theta} \lambda_A \mathbb{1}_A = \sum_{B \in \Theta'} \lambda'_B \mathbb{1}_{B'}.$$

We aim to check that

$$\sum_{A \in \Theta} \lambda_A \mu(A) = \sum_{B \in \Theta'} \lambda'_B \mu(B).$$

Take $C_1, \cdots, C_n \in \mathcal{C}$, pairwise disjoint, such that exch $A \in \Theta \cup \Theta'$ can be written as the union of some sets among $\{C_1, \cdots, C_n\}$.

$$f = \sum_{i=1}^{n} \left(\sum_{\substack{A \in \Theta \\ C_i \cap A \neq \varnothing}} \lambda_A \right) \mathbb{1}_{C_i} = \sum_{i=1}^{n} \left(\sum_{\substack{B \in \Theta' \\ C_i \cap B \neq \varnothing}} \lambda'_B \right) \mathbb{1}_{C_i}.$$

$$\forall i \in \{1, \cdots, n\}, \quad \sum_{\substack{A \in \Theta \\ C_i \cap A \neq \varnothing}} \lambda_A = \sum_{\substack{B \in \Theta' \\ C_i \cap B \neq \varnothing}} \lambda'_B.$$

$$\begin{aligned}\sum_{A\in\Theta}\lambda_A\mu(A) &= \sum_{A\in\Theta}\lambda_A\sum_{\substack{i\in\{1,\cdots,n\}\\ A\cap C_i\neq\varnothing}}\mu(C_i)\\ &= \sum_{i=1}^n\mu(C_i)\sum_{\substack{A\in\Theta\\ A\cap C_i\neq\varnothing}}\lambda_A\\ &= \sum_{i=1}^n\mu(C_i)\sum_{\substack{B\in\Theta'\\ B\cap C_i\neq\varnothing}}\lambda'_B\\ &= \sum_{B\in\Theta'}\lambda'_B\sum_{\substack{i\in\{1,\cdots,n\}\\ B\cap C_i\neq\varnothing}}\mu(C_i)\\ &= \sum_{B\in\Theta'}\lambda'_B\mu(B).\end{aligned}$$

(2) We take, for any $A\in\mathcal{A}$, $\mu(A)$ as $I(\mathbb{1}_A)$. If A is write as a disjoint union $B_1\cup\cdots\cup B_m$, with $B_i\in\mathcal{A}$.

$$I(\mathbb{1}_A)=I(\sum_{j=1}^m\mathbb{1}_{B_j})=\sum_{j=1}^m I(\mathbb{1}_{B_j}).$$

□

Definition 10.6.5 Let Ω be a set and $\mathcal{C}\subseteq\mathscr{P}(\Omega)$. We say that a mapping $\mu:\mathcal{C}\longrightarrow\mathbb{R}_{\geq0}$ is σ-additive if, for any countable set Θ and any family $(C_i)_{i\in\Theta}$ of pairwise disjoint sets in $\mathcal{C}$, one has

$$\bigcup_{i\in\Theta}C_i\in\mathcal{C}\Rightarrow\mu\left(\bigcup_{i\in\Theta}C_i\right)=\sum_{i\in\Theta}\mu(C_i).$$

$$\left(\sum_{i\in\Theta}\mu(C_i)\right):=\sup_{\substack{\Theta'\subseteq\Theta\\ \Theta'\text{ finite}}}\sum_{i\in\Theta'}\mu(C_i).$$

Proposition 10.6.6 Let Ω be a set and $\mathcal{C}$ be a semiring on Ω.

$$\mathcal{A}:=\{A\in\mathscr{P}(\Omega)\mid\exists n\in\mathbb{N},\exists(A_1,\cdots,A_n)\in\mathcal{C}^n,A=A_1\cup\cdots\cup A_n\}.$$

Let $\mu:\mathcal{C}\longrightarrow\mathbb{R}_{\geq0}$ be an additive mapping that extends in a unique way to an additive mapping. $\mu:\mathcal{A}\longrightarrow\mathbb{R}_{\geq0}$, $S\subseteq\mathbb{R}^\Omega$ be the vector subspace generated by $\mathbb{1}_A$, $A\in\mathcal{C}$. Then the following conditions are equivalent:
(1) $\mu:\mathcal{C}\longrightarrow\mathbb{R}_{\geq0}$ is σ-additive.

(2) $\mu : \mathcal{A} \longrightarrow \mathbb{R}_{\geq 0}$ is σ-additive.
(3) For any decreasing sequence $(A_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ such that $\bigcap\limits_{n\in\mathbb{N}} A_n = \varnothing$,

$$\lim_{n\to+\infty} \mu(A_n) = 0.$$

(4) $I : S \longrightarrow \mathbb{R}$, ($\mathbb{R}$-linear mapping) that tends $\mathbb{1}_A$, $A \in \mathcal{C}$ to $\mu(A)$ is an integral operator.

Proof
(1)⇒(2) Let $(A_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ be pairwise disjoint sequence.

$$\forall A_n = \bigcup_{j=0}^{d_n} C_{n,j},\ C_{n,j} \in \mathcal{C},\ (C_{n,j})_{j=0}^{d_n} \text{ pairwise disjoint.}$$

$$A := \bigcup_{n\in\mathbb{N}} A_n,\ A = B_1 \cup \cdots \cup B_n \text{ disjoint union } B_i \in \mathcal{C}.$$

$$B_i = B_i \cap A = \bigcup_{n\in\mathbb{N}} \bigcup_{j=0}^{d_n} (B_i \cap C_{n,j}).$$

Since μ is σ-additive on $\mathcal{C}$,

$$\mu(B_i) = \sum_{n\in\mathbb{N}} \sum_{j=0}^{d_n} \mu(B_i \cap C_{n,j}) = \lim_{N\to+\infty} \sum_{n=0}^{N} \sum_{j=0}^{d_n} \mu(B_i \cap C_{n,j}).$$

$$\begin{aligned}
\mu(A) = \sum_{i=1}^{m} \mu(B_i) &= \lim_{N\to+\infty} \sum_{i=1}^{N} \sum_{n=0}^{N} \sum_{j=0}^{d_n} \mu(B_i \cap C_{n,j}) \\
&= \lim_{N\to+\infty} \sum_{n=0}^{N} \sum_{i=0}^{m} \sum_{j=0}^{d_n} \mu(B_i \cap C_{n,j}) \\
&= \lim_{N\to+\infty} \mu(A_n) = \sum_{n\in\mathbb{N}} \mu(A_n).
\end{aligned}$$

(2)⇒(3) Let $n \in \mathbb{N}$, $B_n = A_n \backslash A_{n+1} \in \mathcal{A}$. Then $\mu(A_n) = \mu(B_n) + \mu(A_{n+1})$.

$$\mu(A_0) = \mu(B_0) + \mu(B_1) + \cdots + \mu(B_{N-1}) + \mu(A_N),\ \forall N \in \mathbb{N}.$$

A_0 is a disjoint union of $(B_n)_{n\in\mathbb{N}}$,

$$\mu(A_0) = \lim_{N\to+\infty} \sum_{n=0}^{N-1} \mu(B_n) = \lim_{N\to+\infty} (\mu(A_0) - \mu(A_N)).$$

So

$$\lim_{N\to+\infty} \mu(A_N) = 0.$$

(3)⇒(4)Let $(f_n)_{n\in\mathbb{N}}$ be a decreasing in S converging pointwise to 0 (as a mapping).

$$B := \{\omega \in \Omega \mid f_0(\omega) > 0\} \in \mathcal{A},\ M = \max\{f_0(\omega) \mid \omega \in \Omega\}.$$

For any $\varepsilon > 0$, $n \in \mathbb{N}$, $A_n^\varepsilon := \{\omega \in \Omega \mid f_n(\omega) \geq \varepsilon\} \in \mathcal{A}$, $(A_n^\varepsilon)_{n\in\mathbb{N}}$ decreasing in $\mathcal{A}$. Since $\lim\limits_{n\to+\infty} f_n = 0$, $\bigcap\limits_{n\in\mathbb{N}} A_n^\varepsilon = \varnothing$.

$$\forall n \in \mathbb{N},\ 0 \leq f_n \leq \varepsilon \mathbb{1}_B + M\mathbb{1}_{A_n^\varepsilon},$$

$$\forall n \in \mathbb{N}, 0 \leq I(f_n) \leq \varepsilon I(\mathbb{1}_B) + M\mu(A_n^\varepsilon).$$

(4)⇒(1) Let $(C_n)_{n\in\mathbb{N}}$ pairwise disjoint in $\mathcal{C}$, $A := \bigcup_{n\in\mathbb{N}} C_n \in \mathcal{C}$. $\forall n \in \mathbb{N}$, $f_n := \sum_{k=0}^n \mathbb{1}_{C_k}$ (increasing) converging pointwise to $\mathbb{1}_A$. $(\mathbb{1}_A - f_n)_{n\in\mathbb{N}} \to 0$ pointwise.

$$\mu(A) := I(\mathbb{1}_A) = \lim_{n\to+\infty} I(f_n) = \lim_{n\to+\infty} \mu(C_k) = \sum_{k\in\mathbb{N}} \mu(C_k).$$

□

Proposition 10.6.7 Let $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ be an increasing right continuous mapping. Let $\mathcal{C} \subseteq \mathscr{P}(\mathbb{R})$ consisting of $]a, b]$, $a, b \in \mathbb{R}$, $a < b$, $S \subseteq \mathbb{R}^\mathbb{R}$ vector subspace generated by $\mathbb{1}_A$, $A \in \mathcal{C}$. $I_\varphi : S \longrightarrow \mathbb{R}$ be $\mathbb{R}$-linear mapping, $I_\varphi(\mathbb{1}_{]a,b]}) := \varphi(b) - \varphi(a)$. $\mathcal{A} \subseteq \mathscr{P}(\mathbb{R})$.

$$\mathcal{A} := \{A \mid A = C_1 \cup \cdots \cup C_n,\ C_1, \cdots, C_n \in \mathcal{C}\}.$$

(1) Let $\varepsilon > 0$, $A \in \mathcal{A}$, $A \neq \varnothing$. There exists $B \in \mathcal{A}$, $\varnothing \neq \overline{B} \subseteq A$ and $I_\varphi(\mathbb{1}_A) - I_\varphi(\mathbb{1}_B) \leq \varepsilon$.
(2) I_φ is an integral operator.
I_φ is typically called Steltjes integral.

Proof
(1)
(A) $A \in \mathcal{C}$, $A =]a, b]$. There exists $a' \in]a, b[$ such that $\varphi(a') - \varphi(a) \leq \varepsilon$ (by right continuity of φ). $B :=]a', b]$, $\overline{B} = [a', b] \subseteq]a, b] = A$. $I_\varphi(\mathbb{1}_B) = \varphi(b) - \varphi(a)$. So $I_\varphi(\mathbb{1}_A) - I_\varphi(\mathbb{1}_B) \leq \varepsilon$.
(B) $A \in \mathcal{A}$, $A = A_1 \cup \cdots \cup A_n$, $\{A_1, \cdots, A_n\} \subseteq \mathcal{C}$, for any $i \in \{1, \cdots, n\}$,

$B_i \in \mathcal{C}$, $\varnothing \neq \overline{B}_i \subseteq A_i$, $I_\varphi(\mathbb{1}_{A_i}) - I_\varphi(\mathbb{1}_{B_i}) \leq \frac{\varepsilon}{n}$. $B := \bigcup B_i$.

$$I_\varphi(\mathbb{1}_A) - I_\varphi(\mathbb{1}_B) = \sum_{i=1}^{n} \left(I_\varphi(\mathbb{1}_{A_i}) - I_\varphi(\mathbb{1}_{B_i})\right) \leq \varepsilon.$$

(2) Let $(A_n)_{n\in\mathbb{N}}$ be a decreasing sequence in $\mathcal{A}$ such that $\bigcap_{n\in\mathbb{N}} A_n = \varnothing$, fix $\varepsilon > 0$, $n \in \mathbb{N}$, $B_n \in \mathcal{A}$, $\overline{B}_n \subseteq A_n$ connected and non-empty, $I_\varphi(\mathbb{1}_{A_n}) - I_\varphi(\mathbb{1}_{B_n}) \leq \frac{\varepsilon}{2^n}$. $\forall n \in \mathbb{N}$, $C_n := B_0 \cap \cdots \cap B_n \subseteq \overline{B_0} \cap \cdots \cap \overline{B}_n$ ($\varnothing = \bigcap_{n\in\mathbb{N}} A_n \supseteq \bigcap_{n\in\mathbb{N}} \overline{B}_n = \varnothing$). So

$$\bigcap_{n=0}^{N} = \overline{B}_n = \varnothing.$$

$B_n \backslash C_n = B_n \backslash (B_n \cap C_{n-1}) = B_n \backslash C_{n-1} \subseteq A_n \backslash C_{n-1} \subseteq A_{n-1} \backslash C_{n-1}$.

$$I_\varphi(\mathbb{1}_{A_n\backslash C_n}) = I_\varphi\left(\mathbb{1}_{B_n\backslash C_n}\right) + I_\varphi\left(\mathbb{1}_{A_n\backslash B_n}\right) \leq I_\varphi\left(\mathbb{1}_{A_{n-1}\backslash C_{n-1}}\right) + \frac{\varepsilon}{2^n}.$$

$\forall n \in \mathbb{N}$, $I_\varphi(\mathbb{1}_{A_n}) \leq \frac{\varepsilon}{2^n} + \frac{\varepsilon}{2^{n-1}} + \cdots + \frac{\varepsilon}{2} \leq \varepsilon$. So $\lim_{n\to+\infty} I_\varphi(\mathbb{1}_{A_n}) = 0$. □

10.7 Measurable Space

Definition 10.7.1 Let Ω be a set. σ-**algebra** on Ω is any $\mathcal{A} \subseteq \mathscr{P}(\Omega)$ such that
(1) For any I countable, $(A_i)_{i\in I} \in \mathcal{A}^I$, $\bigcup_{i\in I} A_i \in \mathcal{A}$.
(2) $A \in \mathcal{A}$, then $\Omega \backslash A \in \mathcal{A}$.
Note that (1) implies that $\varnothing \in \mathcal{A}$, in addition of (2), $\Omega \in \mathcal{A}$.

$$\bigcap_{i\in I} A_i = \Omega \backslash \left(\bigcup_{i\in I} (\Omega \backslash A_i)\right) \in \mathcal{A}.$$

If $\mathcal{A}$ is a σ-algebra on set Ω, then we call $(\Omega, \mathcal{A})$ a **measurable space**.

Proposition 10.7.2 Let I be a countable set, Ω be a set, $J \neq \varnothing$, $(\mathcal{A}_j)_{j\in J}$, $\mathcal{A}_j$ is a σ-algebra on Ω, then $\mathcal{A} := \bigcap_{j\in J} \mathcal{A}_j$ is a σ-algebra on Ω.

Proof Let $(A_i)_{i\in I} \in \mathcal{A}^I$. $\forall j \in J$, $(A_i)_{i\in I} \in \mathcal{A}_j^I$, so $\bigcup_{i\in I} A_i \in \mathcal{A}_j$ and hence

$$\bigcup_{i\in I} A_i \in \bigcap_{j\in J} \mathcal{A}_j = \mathcal{A}.$$

$A \in \mathcal{A}$, so $\forall j \in J$, $A \in \mathcal{A}_j$, so $\Omega \backslash A \in A_j$ and hence

$$\Omega \backslash A \in \bigcap_{j \in J} \mathcal{A}_j = \mathcal{A}.$$

□

Example 10.7.3 Let Ω be a set. Then $\mathscr{P}(\Omega)$ is a σ-algebra on Ω. Moreover, if $\mathcal{C}$ is a subset of $\mathscr{P}(\Omega)$, we denote by $\sigma(\mathcal{C})$ the intersection of all σ-algebras containing $\mathcal{C}$. It is a σ-algebra on Ω, which is the smallest σ-algebra containing $\mathcal{C}$. We call it the σ-algebra generated by $\mathcal{C}$.

Example 10.7.4 Let $(X, \mathscr{T})$ be a topological space. The σ-algebra generated by $\mathscr{T}$ is called the **Borel σ-algebra** on $(X, \mathscr{T})$.

Proposition 10.7.5 Let $A \subseteq [-\infty, +\infty]$. We define a binary relation on A: $(x, y) \in A \times A$, $x \sim y$ if and only if there exists an interval J contained in A such that $\{x, y\} \subseteq J$. Then $\sim$ is an equivalent relation on A. Each equivalent class is an interval.

Proof By definition, if $x \sim y$ and $y \sim z$, $(J_1, J_2) \in A^2$, $\{x, y\} \subseteq J_1$, $\{y, z\} \subseteq J_2$, so $\varnothing \neq J_1 \cap J_2 \subseteq A$, $\{x, z\} \subseteq J_1 \cup J_2$, so $x \sim z$.
Let α be an equivalent class and $(x, y) \in \alpha^2$, $x < y$. $\{x, y\} \subseteq J \subseteq A$, hence $[x, y] \subseteq J$. Hence $\forall z \in [x, y]$, $z \in \alpha$. So α is an interval. □

Remark 10.7.6 Equip $[-\infty, +\infty]$ with order topology, $U \subseteq [-\infty, +\infty]$ open. Consider $\sim$ for U, α an equivalent class. For any $x \in \alpha$, there exists a neighborhood of x contained in U, which is an interval. Hence α is a neighborhood of x. In particular, α is open. U is a disjoint union of open intervals. Any open interval in $[-\infty, +\infty]$ contains a rational number. Hence any open $U \subseteq [-\infty, +\infty]$ is a disjoint union of countable open intervals.

Example 10.7.7 We equip the extended real line $]-\infty, +\infty]$ with the order topology $\mathscr{T}$. Then its Borel σ-algebra $\sigma(\mathscr{T})$ is generated by intervals of the form $]-\infty, b]$ for $b \in \mathbb{Q}$. In fact, if we denote by $\mathcal{A}$ the σ-algebra

$$\sigma(\{]-\infty, b] \mid b \in \mathbb{Q}\}),$$

then by definition one has $\mathcal{A} \subseteq \sigma(\mathscr{T})$.

Conversely, for any $x \in \mathbb{R} \cup \{+\infty\}$, one has

$$]-\infty, x] = \bigcup_{b \in \mathbb{Q},\, b<x}]-\infty, b] \in \mathcal{A}.$$

We then deduce that, for any $x \in]-\infty, +\infty]$, one has

$$[x, +\infty[=]-\infty, +\infty] \setminus]-\infty, x] \in \mathcal{A}.$$

Finally, for any $x \in \mathbb{R} \cup \{-\infty\}$, one has

$$[x, +\infty] = \bigcup_{a \in \mathbb{Q},\, a>x} [a, +\infty[\in \mathcal{A}.$$

Moreover, for any $(a, b) \in]-\infty, +\infty]^2$ such that $a < b$, one has

$$[a, b] =]-\infty, b] \cap [a, +\infty].$$

Therefore, all intervals that are open in $]-\infty, +\infty]$ belongs to $\mathcal{A}$. Finally, by Remark 10.7.6, we obtain that any open subset of $]-\infty, +\infty]$ belongs to $\mathcal{A}$.

Definition 10.7.8 Let $(\Omega_1, \mathcal{A}_1)$, $(\Omega_2, \mathcal{A}_2)$ be measurable spaces, $f : \Omega_1 \longrightarrow \Omega_2$ be a mapping. We say that f is $\mathcal{A}_1$ measurable, if

$$\forall A \in \mathcal{A}_2,\ f^{-1}(A) \in \mathcal{A}_1.$$

Remark 10.7.9 Let $(\Omega, \mathcal{A})$ be a measurable space, $A \in \mathcal{A}$, then $\mathbb{1}_A$ is measurable.

Proposition 10.7.10 Let $(\Omega_1, \mathcal{A}_1)$, $(\Omega_2, \mathcal{A}_2)$, $(\Omega_2, \mathcal{A}_3)$ be measurable spaces. $f : \Omega_1 \longrightarrow \Omega_2$, $g : \Omega_2 \longrightarrow \Omega_3$ be mappings. If f and g are measurable, then $f \circ g$ is measurable.

Proof Let $A \in \mathcal{A}_3$, $g^{-1}(A) \in \mathcal{A}_2$, $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A)) \in \mathcal{A}_1$. □

Proposition 10.7.11 Let $(\Omega, \mathcal{A})$ be a measurable space, X be a set, $\mathcal{C} \subseteq \mathscr{P}(X)$, $f : \Omega \longrightarrow X$ be a mapping. If $\forall A \in \mathcal{C}$, $f^{-1}(A) \in \mathcal{A}$, then f is measurable for X considered $\sigma(\mathcal{C})$. In particular, continuous mappings are measurable (Borel σ-algebras)

Proof Let
$$\mathcal{A}' := \{A \in \mathscr{P}(X) \mid f^{-1}(A) \in \mathcal{A}\}, \mathcal{C} \subseteq \mathcal{A}'.$$
We will show that $\mathcal{A}'$ is a σ-algebra.

Let I be a countable set, $(A_i)_{i\in I} \in \mathcal{A}'^I$,
$$f^{-1}\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} f^{-1}(A_i) \in \mathcal{A}.$$
So, $\bigcup\limits_{i\in I} A_i \in \mathcal{A}'$. For $A \in \mathcal{A}'$,
$$f^{-1}(X\backslash A) = \Omega\backslash f^{-1}(A) \in \mathcal{A}.$$
So, $X\backslash A \in \mathcal{A}'$. Thus $\mathcal{A}'$ is a σ-algebra, $\sigma(\mathcal{C}) \subseteq \mathcal{A}'$. $\square$

Definition 10.7.12 Let Ω be a set, $(E_i, \mathcal{E}_i)_{i\in\Theta}$ be measurable spaces.

For any $i \in \Theta$, fix $f_i : \Omega \longrightarrow E_i$, $f = (f_i)_{i\in\Theta}$.
$$\sigma(f) := \sigma\left(\bigcup_{i\in\Theta} \{f_i^{-1}(A_i),\ A_i \in \mathcal{E}_i\}\right).$$
It is the smallest σ-algebra make all f_i measurable. If $\Omega = \prod\limits_{i\in\Theta} E_i$, $f_i = \pi_i$ be the projection mapping, then $\sigma(f)$ is called the product σ-algebra of $(\mathcal{E}_i)_{i\in\Theta}$, denoted as
$$\bigotimes_{i\in\Theta} \mathcal{E}_i := \sigma(f),\ f = (f_i)_{i\in\Theta} = (\pi_i)_{i\in\Theta}.$$

Proposition 10.7.13 Let X be a set, $(E_i, \mathcal{E}_i)_{i\in\Theta}$ be measurable spaces, $f_i : X \longrightarrow E_i$ be mappings. $(X, \sigma(f))$, $(\Omega, \mathcal{A})$ measurable spaces, $g : \Omega \longrightarrow X$ a mapping. g is measurable if and only if $f_i \circ g$ is measurable for any $i \in \Theta$.

Proof f_i is measurable by definition. If g is measurable, then $f_i \circ g$ is measurable. Conversely, if $f_i \circ g$ is measurable, $A_i \in \mathcal{E}_i$,
$$g^{-1}\left(f_i^{-1}(A_i)\right) = (f_i \circ g)^{-1}(A_i) \in \mathcal{A}.$$
So g is measurable. $\square$

Remark 10.7.14 Let $(X_1, \mathscr{T}_1), \cdots, (X_n, \mathscr{T}_n)$ be topological spaces, $X = X_1 \times \cdots \times X_n$. The product topology on X was generated by $U_1 \times \cdots \times U_n$ with $U_i \in \mathscr{T}_i$, $X \to X_i$ continuous and measurable, $i \in \{1, \cdots, n\}$. The Borel σ-algebra contains the product of σ-algebra $\bigotimes_{i\in\Theta}(\mathscr{T}_i)$. They are equal if for all $i \in \Theta$, $\mathscr{T}_i$ admits a countable basis $\mathcal{B}_i$. In this case, $\{U_1 \times \cdots \times U_n \mid (U_1, \cdots, U_n) \in \mathcal{B}_1 \times \cdots \times \mathcal{B}_n\}$ generates the product topology, so any open set of X belongs to $\bigotimes_{i\in\Theta} \sigma(\mathscr{T}_i)$ or $\mathbb{R}$, $\mathcal{B} = \{]p, q[\mid (p, q) \in \mathbb{Q}^2, p < q\}$. The product σ-algebra on $\mathbb{R}^n$ equals the Borel σ-algebra on $\mathbb{R}^n$. If $(\Omega, \mathcal{A})$ is a measurable space, $f, g : \Omega \longrightarrow \mathbb{R}^n$ measurable, then $f + g$ and $f \cdot g$ are measurable.

Notation 10.7.15 Let Ω be a set, $f : \Omega \longrightarrow [-\infty, +\infty]$ be a mapping, P be a condition on $[-\infty, +\infty]$, $\{P(f)\}$ denotes the set

$$f^{-1}(\{t \in [-\infty, +\infty] \mid P(t)\}) = \{\omega \in \Omega \mid P(f(\omega))\}.$$

For example, $\{f > 0\} = \{\omega \in \Omega \mid f(\omega) > 0\}$.

Remark 10.7.16 Let Ω be a set, $f : \Omega \longrightarrow \mathbb{R}_{\geq 0}$ be a mapping, $\forall n \in \mathbb{N}$,

$$f_n := \sum_{k=1}^{n\cdot 2^n - 1} \frac{k}{2^n} \mathbb{1}_{\{\frac{k}{2^n} \leq f \leq \frac{k+1}{2^n}\}} + \mathbb{1}_{\{f \geq n\}}.$$

Theorem 10.7.17 Let $(\Omega, \mathcal{A})$ be a measurable space, $(X, \mathscr{T})$ be a topological space such that $\mathscr{T}$ is given by metric d. $(f_n)_{n\in\mathbb{N}}$ is measurable. If $(f_n)_{n\in\mathbb{N}}$ converges pointwise to some $f : \Omega \longrightarrow X$, then f is measurable.

Proof Let $Y \subseteq X$ be a closed subset, $\mathrm{d}(\cdot, Y) : X \longrightarrow \mathbb{R}_{\geq 0}$ continuous (Lipschitzian).

$$\begin{aligned} f^{-1}(Y) :=& \{\omega \in \Omega \mid \lim_{n\to+\infty} \mathrm{d}(f_n(\omega), Y) = 0\} \\ =& \{\omega \in \Omega \mid \limsup_{n\to+\infty} \mathrm{d}(f_n(\omega), Y) = 0\} \\ =& \bigcap_{m\in\mathbb{N}_{\geq 1}} \bigcup_{N\in\mathbb{N}} \bigcap_{n\in\mathbb{N}_{\geq N}} \{\omega \in \Omega \mid \mathrm{d}(f_n(\omega), Y) < m^{-1}\}. \end{aligned}$$

Since f_n is measurable and $\mathrm{d}(\cdot, Y)$ is continuous (so measurable).

$$\{\omega \in \Omega \mid \limsup_{n\to+\infty} \mathrm{d}(f_n(\omega, Y)) = 0\} \in \mathcal{A}.$$

□

10.8 Monotone Class Theorem

Lemma 10.8.1 Let $n \in \mathbb{N}_{\geq 1}$, and

$$\begin{array}{rccc} \varphi_n : & \mathbb{R}_{\geq 0} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x^n \end{array}$$

be a mapping.
(1) φ_n is convex.
(2) $\forall x \in \mathbb{R}_{\geq 0}$,

$$\varphi_n(x) = x^n = \sup_{a \in \mathbb{Q}_{\geq 0}} \max\{na^{n-1}x - (n-1)a^n, 0\}.$$

Proof
(1) $\varphi''(x) = n(n-1)x^{n-2} \geq 0$.
(2) $\forall a \geq 0$,

$$\varphi_n(x) \geq \varphi_n(a) + \varphi_n'(a)(x-a) = a^n + na^{n-1}(x-a).$$

Since $\mathbb{Q}$ is dense in $\mathbb{R}$ and φ_n is continuous, so (2) holds. □

Definition 10.8.2 Let Ω be a set and $\mathcal{H}$ be a family of bounded mappings from Ω to $\mathbb{R}_{\geq 0}$. We say that $\mathcal{H}$ is a **λ-family**[a] if the following conditions are satisfied:
(1) $\mathbb{1}_\Omega \in \mathcal{H}$.
(2) If $(f, g) \in \mathcal{H} \times \mathcal{H}$, $(a, b) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$, then $af + bg \in \mathcal{H}$.
(3) If $(f, g) \in \mathcal{H} \times \mathcal{H}$, $f \leq g$, then $g - f \in \mathcal{H}$.
(4) If $(f_n)_{n \in \mathbb{N}}$ is an increasing and uniformly bounded sequence in $\mathcal{H}$, then

$$\lim_{n \to +\infty} f_n \in \mathcal{H}.$$

[a] The following monotone class theorem is named by weaker definition: monotone class (of functions) that do not need the condition of non-negative.

Remark 10.8.3 Let $(f_i)_{i \in \Theta}$ is a countable and uniformly bounded family in $\mathcal{H}$. Assume that $\forall (f, g) \in \mathcal{H}$, $\sup\{f, g\} \in \mathcal{H}$, then

$$\sup_{i \in \Theta} f_i \in \mathcal{H}.$$

In fact, we assume that $\Theta = \mathbb{N}$. For any $n \in \mathbb{N}$, let $g_n = \sup\{f_0, \cdots, f_n\}$, then $(g_n)_{n \in \mathbb{N}} \in \mathcal{H}^{\mathbb{N}}$ is increasing, uniformly bounded, and converges to $\sup_{i \in \Theta} f_i \in \mathcal{H}$.

Theorem 10.8.4 (Monotone class theorem) Let Ω be a set. Let $\mathcal{H}$ be a λ-family on Ω. If $\forall (f,g) \in \mathcal{H}\times\mathcal{H}$, $\inf\{f,g\} \in \mathcal{H}$. Then any bounded $\sigma(\mathcal{H})$-measurable mapping from Ω to $\mathbb{R}_{\geq 0}$ belongs to $\mathcal{H}$.

Proof $\forall (f,g) \in \mathcal{H}\times\mathcal{H}$, $\sup\{f,g\} = f+g-\inf\{f,g\} \in \mathcal{H}$. Moreover, for any $a \in \mathbb{R}_{\geq 0}$, $\forall f \in \mathcal{H}$,

$$\sup\{f - a\mathbb{1}_\Omega, 0\} = \sup\{f, a\mathbb{1}_\Omega\} - a\mathbb{1}_\Omega \in \mathcal{H}.$$

We then deduce by the lemma that $\forall n \in \mathbb{N}_{\geq 1}$, $f^n \in \mathcal{H}$. In fact, by the lemma

$$f^n = \sup_{a \in \mathbb{Q}_{>0}} \big(\sup\{na^{n-1}f - (n-1)a^n\mathbb{1}_\Omega, 0\}\big) \in \mathcal{H}.$$

Let $\mathcal{A} = \{A \in \mathscr{P}(\Omega) \mid \mathbb{1}_A \in \mathcal{H}\}$.
(1) If $A \in \mathcal{A}$, $\mathbb{1}_{\Omega\setminus A} = \mathbb{1}_\Omega - \mathbb{1}_A \in \mathcal{H}$.
(2) If $(A_i)_{i\in\Theta} \in \mathcal{A}^\Theta$, with Θ countable, $A = \bigcup\limits_{i\in\Theta} A_i$. $\mathbb{1}_A = \sup_{i\in\Theta}\mathbb{1}_{A_i} \in \mathcal{H}$.
Therefore $\mathcal{A}$ is a σ-algebra contained in $\sigma(\mathcal{H})$.

Let $f \in \mathcal{H}$, $t > 0$. One has $\inf\{\mathbb{1}_\Omega, t^{-1}f\} \in \mathcal{H}$. So $\forall n \in \mathbb{N}_{\geq 1}$, $\inf\{\mathbb{1}_\Omega, t^{-1}f\}^n \in \mathcal{H}$.

$$\big(\mathbb{1}_\Omega - \inf\{\mathbb{1}_\Omega, t^{-1}f\}^n\big)_{n\in\mathbb{N}}$$

is increasing and converges to $\mathbb{1}_{\{f<t\}}$. So $\mathbb{1}_{\{f<t\}} \in \mathcal{H}$, $\{f<t\} \in \mathcal{A}$, so f is $\mathcal{A}$-measurable. (In fact, $\sigma(\mathcal{H}) = \mathcal{A}$. By definition, $\sigma(\mathcal{H})$ is the smallest set making all the mappings in $\mathcal{H}$ measurable, so $\sigma(\mathcal{H})$ is contained in $\mathcal{A}$.) □

Theorem 10.8.5 Let Ω be a set and $\mathcal{H}$ be a λ-family on Ω. Let $\mathcal{H}_0 \subseteq \mathcal{H}$. Suppose that $\forall (f,g) \in \mathcal{H}_0\times\mathcal{H}_0$, $fg \in \mathcal{H}_0$. Then any bounded $\sigma(\mathcal{H}_0)$-measurable mapping $\Omega \longrightarrow \mathbb{R}_{\geq 0}$ belongs to $\mathcal{H}$.

Proof We may assume that $\mathcal{H}$ is the smallest λ-family containing $\mathcal{H}_0$ (by taking the intersection of all λ-families containing $\mathcal{H}_0$). Let $\mathcal{H}_1 = \{f \in \mathcal{H} \mid \forall g \in \mathcal{H}_0, fg \in \mathcal{H}\} \supseteq \mathcal{H}_0$. Moreover, $\mathcal{H}_1$ is a λ-family, so $\mathcal{H}_1 = H$. Let $\mathcal{H}_2 = \{f \in \mathcal{H} \mid \forall g \in \mathcal{H}, fg \in \mathcal{H}\} \supseteq \mathcal{H}_0$. $\mathcal{H}_2$ is a λ-family, so $\mathcal{H}_2 = \mathcal{H}$. Therefore, $\mathcal{H}$ is stable by multiplication.
Let $(f,g) \in \mathcal{H}\times\mathcal{H}$, $|f-g| \leq 1$.

$$(f-g)^2 = f^2 + g^2 - 2fg \in \mathcal{H}.$$

$(z \longmapsto z - \frac{1}{2}z^2)$ is increasing on $[0,1]$. Let $(f_n)_{n\in\mathbb{N}}$ be the sequence of mappings

defined as $f_0 = 0$, $f_{n+1} = f_n + \frac{1}{2}((f-g)^2 - f_n^2)$. We prove by induction that $f_n \in \mathcal{H}$ and $f_n \leq |f-g|$.
$n = 0$, $f_0 = 0 \in \mathcal{H}$ and $f_0 \leq |f-g|$. Suppose $f_n \in \mathcal{H}$, $f_n \leq |f-g|$. Then $f_{n+1} \in \mathcal{H}$, $f_{n+1} \geq f_n$.

$$f_{n+1} = \varphi(f_n) + \frac{1}{2}(f-g)^2 \leq \varphi(|f-g|) + \frac{1}{2}(f-g)^2 = |f-g|.$$

$(f_n)_{n\in\mathbb{N}}$ converges to $|f-g|$. So $|f-g| \in \mathcal{H}$. $\inf\{f, g\} = \frac{1}{2}(f+g-|f-g|) \in \mathcal{H}$. So any bounded $\sigma(\mathcal{H})$-measurable mappings $\Omega \longrightarrow \mathbb{R}_{\geq 0}$ belongs to $\mathcal{H}$. $\square$

Theorem 10.8.6 Let Ω be a set, $\mathcal{L}$ be a Riesz space on Ω. We assume that $\forall (f, g) \in \mathcal{L} \times \mathcal{L}^{\uparrow}$, $\inf\{f, g\} \in \mathcal{L}$. Then the following statements hold.
(1) $\forall (f, g) \in \mathcal{L}^{\uparrow} \times \mathcal{L}^{\uparrow}$, $f \leq g$ and $g - f$ is well defined, then $g - f \in \mathcal{L}^{\uparrow}$.
(2) Let $\mathcal{A} = \{A \in \mathscr{P}(\Omega) \mid \mathbb{1}_A \in \mathcal{L}^{\uparrow}\}$.

1. If $(A, B) \in \mathcal{A}^2$, $B \backslash A \in \mathcal{A}$.
2. If Θ is countable and $(A_i)_{i\in\Theta} \in \mathcal{A}^{\Theta}$, then $\bigcup\limits_{i\in\Theta} A_i \in \mathcal{A}$.

(3) Let $f \in \mathcal{L}^{\uparrow}$, $f \geq 0$. Suppose that

$$B := \{\omega \in \Omega \mid f(\omega) > 0\} \in \mathcal{A}.$$

Then, $\forall t > 0$, $\{\omega \in \Omega \mid 0 < f(\omega) < t\} \in \mathcal{A}$.
(4) Let $f : \Omega \longrightarrow [0, +\infty]$ such that

$$B := \{\omega \in \Omega \mid f(\omega) > 0\} \in \mathcal{A}.$$

If $\forall t > 0$, $A_t := \{\omega \in \Omega \mid 0 < f(\omega) < t\} \in \mathcal{A}$, then $f \in \mathcal{L}^{\uparrow}$.

Proof
(1) Let $(g_n)_{n\in\mathbb{N}} \in \mathcal{L}^{\mathbb{N}}$ be an increasing sequence such that $g = \lim\limits_{n\to+\infty} g_n$. $\forall n \in \mathbb{N}$, $\inf\{g_n, f\} \in \mathcal{L}$ and $\lim\limits_{n\to+\infty} \inf\{g_n, f\} = f$,

$$(g_n - \inf\{g_n, f\})_{n\in\mathbb{N}} = (\sup\{g_n - f, 0\})_{n\in\mathbb{N}}$$

is increasing and converges to $g - f$. So $(g - f) \in \mathcal{L}^{\uparrow}$.
(2)

1.
$$\mathbb{1}_{A\cap B} = \inf\{\mathbb{1}_A, \mathbb{1}_B\} \in \mathcal{L}^{\uparrow}.$$
$$\mathbb{1}_{B\setminus A} = \mathbb{1}_B - \mathbb{1}_{A\cap B} \in \mathcal{L}^{\uparrow}(\text{by (1)}).$$

2. If $(A, B) \in \mathcal{A}^2$, then
$$\mathbb{1}_{A\cup B} = \sup\{\mathbb{1}_A, \mathbb{1}_B\} \in \mathcal{L}^{\uparrow}.$$

So $A \cup B \in \mathcal{A}$. The case where Θ is finite is true. We assume that $\Theta = \mathbb{N}$.
$$\mathbb{1}_{\bigcup_{n\in\mathbb{N}} A_n} = \lim_{n\to+\infty} (\sup\{\mathbb{1}_{A_0}, \cdots, \mathbb{1}_{A_n}\}) \in \mathcal{L}^{\uparrow}.$$

(3)
$$\sup\{f - a\mathbb{1}_B, 0\} = \sup\{f, a\mathbb{1}_B\} - a\mathbb{1}_B \in \mathcal{L}^{\uparrow}.$$
$$f^n = \sup_{a\in\mathbb{Q}_{>0}} \sup\{na^{n-1}f - (n-1)a^n\mathbb{1}_B, 0\} \in \mathcal{L}^{\uparrow}.$$
$$\inf\{t^{-1}f, \mathbb{1}_B\}^n \in \mathcal{L}^{\uparrow}, \forall t > 0,\ \forall n \in \mathbb{N}_{\geq 1}.$$
$(\mathbb{1}_B - \inf\{t^{-1}f, \mathbb{1}_B\}^n)_{n\in\mathbb{N}_{\geq 1}}$ is increasing and converges to $\mathbb{1}_{\{0<f<t\}}$. So, $\{0 < f < t\} \in \mathcal{A}$.
(4) Let
$$f_n = \sum_{k=1}^{n2^n-1} \frac{k}{2^n}\mathbb{1}_{\{\frac{k}{2^n}\leq f<\frac{k+1}{2^n}\}} + n\mathbb{1}_{\{f\geq n\}} = \sum_{k=1}^{n2^n-1} \frac{k}{2^n}\mathbb{1}_{A_{\frac{k+1}{2}}\setminus A_{\frac{k}{2^n}}} + n\mathbb{1}_{B\setminus A} \in \mathcal{L}^{\uparrow}.$$
Since f is the limit of $(f_n)_{n\in\mathbb{N}}$, we obtain $f \in \mathcal{L}^{\uparrow}$. □

Remark 10.8.7 Let Ω be a set, S be a Riesz space on Ω. $I : S \longrightarrow \mathbb{R}$ be an integral operator. If $(f, g) \in \mathcal{L}^1(I) \times \mathcal{L}^1(I)^{\uparrow}$, then
$$\inf\{f, g\} \in \mathcal{L}^1(I)^{\uparrow} \text{ and } I(\inf\{f, g\}) \leq I(f) < +\infty.$$
So, $\inf\{f, g\} \in \mathcal{L}^1(I)$ by Beppo Levi's theorem.

10.9 Measure Space

Definition 10.9.1 Let $(\Omega, \mathcal{A})$ be a measurable space. We call **measure** on $(\Omega, \mathcal{A})$ any mapping $\mu : \mathcal{A} \longrightarrow [0, +\infty]$ such that
(1) $\mu(\varnothing) = 0$.
(2) μ is σ-additive.
$(\Omega, \mathcal{A}, \mu)$ is called a **measure space**.

Remark 10.9.2 Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Let $(A_n)_{n\in\mathbb{N}}$ be an increasing sequence in $\mathcal{A}$ and $A = \bigcup\limits_{n\in\mathbb{N}} A_n$. Then

$$\mu(A) = \lim_{n\to+\infty} \mu(A_n).$$

In fact, if we let $B_n = A_n \backslash A_{n-1}$, $(n \geq 1)$, $B_0, \cdots, B_n$ are pairwise disjoint and $B_0 = A_0$.

$$A_n = B_0 \cup \cdots \cup B_n, \mu(A_n) = \mu(B_0) + \cdots + \mu(B_n).$$

So,

$$\lim_{n\to+\infty} \mu(A_n) = \sum_{n\in\mathbb{N}} \mu(B_n) = \mu\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \mu(A).$$

Proposition 10.9.3 Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Let

$$\mathcal{C} := \{A \in \mathcal{A} \mid \mu(A) < +\infty\}.$$

Then $\mathcal{C}$ is a semiring and $\mu|_{\mathcal{C}}$ is σ-additive.

Proof $\mu(\varnothing) = 0 \Rightarrow \varnothing \in \mathcal{C}$. If $(A, B) \in \mathcal{C} \times \mathcal{C}$, $\mu(A \cap B) \leq \mu(A)$ $(\mu(A) = \mu(A \cap B) + \mu(A \backslash B))$. So $A \cap B \in \mathcal{C}$, $A \backslash B \in \mathcal{C}$. □

Definition 10.9.4 Let $(\Omega, \mathcal{A}, \mu)$ be a measure space,

$$\mathcal{C} = \{A \in \mathcal{A} \mid \mu(A) < +\infty\},$$

$$S = \mathrm{Vect}_{\mathbb{R}}\left(\{\mathbb{1}_A \mid A \in \mathcal{C}\}\right) \subseteq \mathbb{R}^{\Omega},$$

$$I : S \longrightarrow \mathbb{R}, I(\mathbb{1}_A) = \mu(A).$$

We denote by $\mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ the set of $\mathcal{A}$-measurable mappings from Ω to $\mathbb{R}$ such that belongs to $\mathcal{L}^1(I)$. If $f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$, we denote by

$$\int_\Omega f \mathrm{d}\mu \text{ or } \int_\Omega f(\omega)\mu(\mathrm{d}\omega)$$

the valued of $I(f)$. If

$$f : \Omega \longrightarrow \mathbb{R} \cup \{+\infty\}$$

is $\mathcal{A}$-measurable and bounded from below by the same element of $\mathcal{L}^1(\Omega, \mathcal{A}, \mu)$. When $f \notin \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$. By convention, $\int_\Omega f \mathrm{d}\mu := +\infty$.

Proposition 10.9.5 For any $A \in \mathcal{A}$, $\displaystyle\int_\Omega \mathbb{1}_A \mathrm{d}\mu = \mu(A)$.

Proof This is true by definition when $\mu(A) < +\infty$. We assume that $\mu(A) = +\infty$. We reason by contradiction that $\mathbb{1}_A \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$.

We first show that $\forall f \in S$, $f > 0$, one has $\inf\{\mathbb{1}_A, f\} \in S$. If

$$f = \sum_{i=1}^n a_i \mathbb{1}_{B_i},$$

then

$$\inf\{\mathbb{1}_A, f\} = \sum_{i=1}^n \min\{b_i, 1\}\mathbb{1}_{A\cap B_i} \in S.$$

Since $\mathbb{1}_A \in \mathcal{L}^1(\Omega, A, \mu)$, $\exists g \in S^\uparrow$, $\mathbb{1}_A \leq g$. $I(g) < +\infty$. Let $(f_n)_{n\in\mathbb{N}} \in S^\mathbb{N}$ increasing such that $g = \lim_{n\to+\infty} f_n$ (we may assume). Then $(\inf\{f_n, \mathbb{1}_A\})_{n\in\mathbb{N}} \in S^\mathbb{N}$ is increasing and converges to $\mathbb{1}_A$. Let $A_n = \{\omega \in A \mid f_n(\omega) \geq 1\} \in \mathcal{C}$, $(A_n)_{n\in\mathbb{N}}$ is increasing. $A = \bigcup_{n\in\mathbb{N}} A_n$.

$$\mu(A) = \lim_{n\to+\infty} \mu(A_n) \leq \int_\Omega g \mathrm{d}\mu < +\infty.$$

Contradiction. □

Proposition 10.9.6 Let $(\Omega, \mathcal{A}, \mu)$ be a measure space,
(1) Let $f : \Omega \longrightarrow \mathbb{R}_{\geq 0}$ be an $\mathcal{A}$-measurable mapping. If $f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$, then $\forall t \in \mathbb{R}_{\geq 0}$, $\mu(\{\omega \in \Omega \mid f(\omega) \geq t\}) < +\infty$.
(2) Let f, g be measurable mappings from Ω to $\mathbb{R}$ such that $f \leq g$. If $g \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ then $f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$
(3) An $\mathcal{A}$-measurable mapping $f : \Omega \longrightarrow \mathbb{R}$ belongs to $\mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ if and only

if $|f| \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$.

Proof Apply monotone class theorem to $\mathcal{L}^1(I)$.
(1) $\mathbb{1}_{0<f<x} \in \mathcal{L}^1(I)^{\uparrow}$, $\forall x > 0$. If $0 < t < x$,

$$\mathbb{1}_{\{t\leq f<x\}} = \mathbb{1}_{\{0<f<x\}} - \mathbb{1}_{\{0<f<t\}} \in \mathcal{L}^1(I)^{\uparrow}.$$

So,

$$\mathbb{1}_{f\geq t} = \lim_{n\to+\infty} \mathbb{1}_{\{t\leq f<t+n\}} \in \mathcal{L}^1(I)^{\uparrow}.$$

Since $t\mathbb{1}_{t\leq f} \leq f$. So, $t\mu(\{t \leq f\}) = I(t\mathbb{1}_{\{t\leq f\}}) \leq I(f) < +\infty$. So $\mu(\{f \geq t\} < +\infty)$.
(2) By (1), $\forall a > 0$, $\mu(\{g \geq a\}) < +\infty$. So $\forall a > 0$, $\mu(\{f \geq a\}) < +\infty$. So,

$$\mathbb{1}_{\{f>0\}} = \sup_{a\in\mathbb{Q}_{>0}} \{\mathbb{1}_{f\geq a}\} \in \mathcal{L}^1(I)^{\uparrow}.$$

For $0 < a < t$, $\mu(\{a \leq f < t\}) \leq \mu(\{f \geq a\}) < +\infty$. So $\mathbb{1}_{\{0<f<t\}} \in \mathcal{L}^1(I)$. Therefore, (by monotone class theorem) $f \in \mathcal{L}^1(I)^{\uparrow}$. Hence $I(f) \leq I(g) < +\infty$. So $f \in \mathcal{L}^(I)$.
(3) $\Rightarrow$: $|f| = \sup\{f, 0\} = \inf\{f, 0\} \in \mathcal{L}^1(I)$, if $f \in \mathcal{L}^1(I)$.
$\Leftarrow$:

$$0 \leq \sup\{f, 0\} \leq |f|,\ 0 \leq \sup\{-f, 0\} \leq |f|.$$

So $\sup\{f, 0\}, \sup\{-f, 0\} \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$. Hence,

$$f = \sup\{f, 0\} - \sup\{-f, 0\} \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu).$$

$\square$

Notation 10.9.7 Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, $f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$. For any $A \in \mathcal{A}$, $|\mathbb{1}_A f| \leq |f|$. So $\mathbb{1}_A f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$. If $\mu = \mu_\varphi$, where $\varphi : [a, b] \longrightarrow \mathbb{R}$ increasing and right continuous,

$$\int_a^b f\mathrm{d}\mu_\varphi \text{ is written as } \int_a^b f(t)\mathrm{d}(\varphi(t)).$$

If $\varphi(t) = t$, it is also written as

$$\int_a^b f(t)\mathrm{d}t.$$

Definition 10.9.8 Let $(\Omega, \mathcal{A}, \mu)$ be measure space. If there exists $(A_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ such that
(1) $\forall n \in \mathbb{N}$, $\mu(A_n) < +\infty$.
(2) $\Omega = \bigcup_{n\in\mathbb{N}} A_n$, then we say that $(\Omega, \mathcal{A}, \mu)$ is σ-finite.

Remark 10.9.9 If $(\Omega, \mathcal{A}, \mu)$ is σ-finite, then $\mathbb{1}_\Omega \in \mathcal{L}^1(I)^\uparrow$.

Theorem 10.9.10 (Carathéodory) Let Ω be a set, $\mathcal{C}$ be a semiring on Ω, $\mu : \mathcal{C} \longrightarrow \mathbb{R}_{\leq 0}$ be a σ-additive mapping. If there exists an increasing sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ such that
(1) $\forall n \in \mathbb{N}$, $\mu(A_n) < +\infty$.
(2) $\Omega = \bigcup\limits_{n\in\mathbb{N}} A_n$.
Then μ extends in a unique way to a σ-finite measure on $\sigma(\mathcal{C})$.

Proof Let $S = \mathrm{Span}_{\mathbb{R}}\left(\{\mathbb{1}_A \mid A \in \mathcal{C}\}\right)$,

$$\begin{array}{cccc} I: & S & \longrightarrow & \mathbb{R} \\ & \mathbb{1}_A & \longmapsto & \mu(A) \end{array}$$

is an integral operator.

$$\mathcal{A} = \{A \in \mathscr{P}(\Omega) \mid \mathbb{1}_A \in \mathcal{L}^1(I)^\uparrow\}$$

is a σ-algebra since $\mathbb{1}_\Omega \in S^\uparrow \subseteq \mathcal{L}^1(I)^\uparrow$ and $(A \in \mathcal{A}) \longmapsto I(\mathbb{1}_A)$ is a measure extending μ. If ν is a measure on $\sigma(\mathcal{C})$ extending μ. Let $S' = \mathrm{Span}_{\mathbb{R}}\{\mathbb{1}_A \mid A \in \sigma(\mathcal{C}), \nu(A) < +\infty\}$. One has $S \subseteq S' \subseteq \mathcal{L}^1(I)$. By Beppo Levi's theorem, the restriction of I_ν to $S^\uparrow$ and $S^\downarrow$ coincides with I. Therefore, $I = I_\nu$ and hence $\nu = \mu$ on $\sigma(\mathcal{C})$. $\square$

Application
Let $\varphi : I \longrightarrow \mathbb{R}$ be an increasing right continuous mapping, where $I \subseteq \mathbb{R}$ is an interval. Let $\mathcal{C} = \{]a, b] \mid (a, b) \in I^2, a < b\}$. $\mu_\varphi : \mathcal{C} \longrightarrow \mathbb{R}_{\geq 0}$, $]a, b] \longmapsto \varphi(b) - \varphi(a)$ is σ-additive. So it extends to a measure μ_φ on $\sigma(\mathcal{C})$ (Borel σ-algebra on I).

Definition 10.9.11 Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. We say that $A \in \mathscr{P}(\Omega)$ is μ-negligeable if there exists $B \in \mathcal{A}$, $\mu(B) = 0$ and $A \subseteq B$. (If $A \in \mathcal{A}$ and is

μ-negligeable, then $\mu(A) = 0$).

Proposition 10.9.12 If $f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ and $A \in \mathcal{A}$, $\mu(A) = 0$, then $\int_\Omega \mathbb{1}_A f \mathrm{d}\mu = 0$.

Proof We may assume that $f \in S = \mathrm{Span}_{\mathbb{R}}\{\mathbb{1}_B \mid B \in \mathcal{A}, \mu(B) < +\infty\}$. When $f = \mathbb{1}_B$, $B \in \mathcal{A}$, $\mu(B) =< +\infty$.

$$\mathbb{1}_A \cdot f = \mathbb{1}_{A \cap B}, \int_\Omega \mathbb{1}_A f \mathrm{d}\mu = \mu(A \cap B) = 0.$$

($I_A : S \longrightarrow \mathbb{R}$, $f \longmapsto \int_\Omega \mathbb{1}_A f \mathrm{d}\mu$ is a integral operator.) □

Theorem 10.9.13 Let Y be a metric space, $(X, \mathcal{A}, \mu)$ be a measure space. $f : X \times Y \longrightarrow \mathbb{R}$ be a mapping, $g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ and $p \in y$. Assume that
(1) There exists μ-negligeable set $S \in \mathscr{P}(\Omega)$ such that $\forall x \in X \backslash A$, $f(x, \cdot) : Y \longrightarrow \mathbb{R}$, $y \longmapsto f(x, y)$ is continuous at p.
(2) For any $y \in Y$, $f(\cdot, y) : X \longrightarrow \mathbb{R}$, $x \longmapsto f(x, y)$ is $\mathcal{A}$-measurable.
(3) $\forall y \in Y$, $\exists A_y \in \mathscr{P}(\Omega)$ μ-negligeable, such that $\forall x \in X \backslash A_y$, $|f(x, y)| \leq g(x)$. Then $(y \in Y) \longmapsto \int_X f(x, y)\mu(\mathrm{d}x)$ is continuous at p.

Proof Let $(y_n)_{n \in \mathbb{N}}$ be a sequence in Y that converges to p. For any $n \in \mathbb{N}$, let $f_n : X \longrightarrow \mathbb{R}$, $f_n(x) := f(x, y_n)$. Let $B = A \cup \left(\bigcup_{n \in \mathbb{N}} A_{y_n}\right)$, $\mu(B) = 0$. On $X \backslash B$, f_n converges pointwise to $f(\cdot, p)$ and $|f_n(x)| \leq g(x)$. By dominated convergence theorem,

$$\lim_{n \to +\infty} \int_{X \backslash B} f(x, y_n)\mu(\mathrm{d}x) = \int_{X \backslash B} f(x, p)\mu(\mathrm{d}x).$$

□

Corollary 10.9.14 Let $(a, b) \in \mathbb{R}^2$, $a < b$, μ be a Borel measure on $[a, b]$, and $f : [a, b] \longrightarrow \mathbb{R}$ be a Borel measurable mapping. Let $F : [a, b] \longrightarrow \mathbb{R}$,

$$F(x) := \int_a^x f(t)\mu(\mathrm{d}t).$$

If $x_0 \in [a, b]$ is such that $\mu(\{x_0\}) = 0$, then F is continuous at x_0.

Proof By definition,

$$\int_a^x f(t)\mu(\mathrm{d}t) := \int_{]a,x]} f(t)\mu(\mathrm{d}t) = \int_{[a,b]} \mathbb{1}_{]a,x]}(t) f(t)\mu(\mathrm{d}t).$$

For any $x_0 \in [a,b]$. If $t \neq x_0$, then $x \longmapsto \mathbb{1}_{]a,x]}(t)$ is continuous at x_0. Since $|\mathbb{1}_{]a,x]} f| \leq |f|$, we obtain that F is continuous at x_0. □

Theorem 10.9.15 Let $(a,b) \in \mathbb{R}^2$, $a < b$ and $f : [a,b] \longrightarrow \mathbb{R}$ be a Borel measurable mapping which is Lebesgue integrable ($f \in \mathcal{L}^1([a,b])$).
Let $F : [a,b] \longrightarrow \mathbb{R}$, $F(x) := \int_a^x f(t)\mathrm{d}t$.
(1) Let $x_0 \in]a,b[$. If f is continuous at x_0, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.
(2) Suppose that f is continuous and $G : [a,b] \longrightarrow \mathbb{R}$ is a continuous mapping such that

$$\forall x \in]a,b[, G'(x) = f(x).$$

Then,

$$\int_a^b f(t)\mathrm{d}t = G(b) - G(a).$$

Moreover, $\exists \xi \in]a,b[$, $\displaystyle\int_a^b f(t)\mathrm{d}t = f(\xi)(b-a)$.

Proof
(1) For $h < 0$, sufficiently small, let

$$\mu_h = \sup_{t\in[x_0-h,x_0+h]} |f(t) - f(x+0)|.$$

$$|F(x_0+h) - F(x_0) - f(x_0)h| = \left|\int_{x_0}^{x_0+h} (f(t) - f(x_0))\mathrm{d}t\right| \leq \mu_h \cdot h = o(h), h \to 0.$$

Similarly,

$$|F(x_0 - h) - F(x_0) + f(x_0)h| \leq \mu_h \cdot h = o(h), h \to 0.$$

(2) By (1), F is a primitive function of f on $[a,b]$, so $F - G$ is a constant mapping. Hence

$$\int_a^b f(t)\mathrm{d}t = F(b) - F(a) = G(b) - G(a).$$

By mean value theorem, $\exists \xi \in]a,b[$, $F(b) - F(a) = F'(\xi)(b-a)$. That is

$$\int_a^b f(t)\mathrm{d}t = f(\xi)(b-a).$$

□

10.10 Product Measure

We fix $(X, \mathcal{A}_X, \mu_X)$ and $(Y, \mathcal{A}_Y, \mu_Y)$ two σ-finite measure space.
We equip $X \times Y$ the product σ-algebra

$$\mathcal{A}_X \otimes \mathcal{A}_Y = \sigma\left(\{A \times B \mid A \in \mathcal{A}_X,\ B \in \mathcal{A}_Y\}\right).$$

Proposition 10.10.1 Let $f : X \times Y \longrightarrow \mathbb{R}$ be a measurable mapping. For any $x \in X$, $f(x, \cdot) : Y \longrightarrow \mathbb{R}$, $y \longmapsto f(x, y)$ is $\mathcal{A}_Y$ measurable.

Proof
(1) $f = \sup\{f, 0\} - \sup\{-f, 0\}$. We may assume that f is non-negative.
(2)

$$f = \lim_{n \to +\infty} \sum_{k=0}^{n \cdot 2^n - 1} \frac{k}{2^n} \mathbb{1}_{\{\frac{k}{2^n} \leq f \leq \frac{k+1}{2^n}\}} + n \mathbb{1}_{\{f \geq n\}}.$$

We may assume that f is bounded. Let

$$\mathcal{H} = \{f : X \times Y \longrightarrow \mathbb{R}_{\geq 0} \mid f \text{ bounded }, f(x, \cdot) \text{ is } \mathcal{A}_Y\text{-measurable}, \forall x \in X\}.$$

Then $\mathcal{H}$ is a λ-family.
If $f : X \times Y \longrightarrow \mathbb{R}_{\geq 0}$ is of the form $f(x, y) = f_1(x) f_2(y)$, where f_1 and f_2 are non-negative, measurable and bounded, $f \in \mathcal{H}$. Let $\mathcal{C}$ be the set of such mappings. Then $\mathcal{C} \subseteq \mathcal{H}$ and $\mathcal{C}$ is stable by multiplication. By monotone class theorem $\mathcal{H}$ contains all bounded non-negative $\mathcal{A}_X \otimes \mathcal{A}_Y = \sigma(\mathcal{C})$-measurable functions. (if $(A, B) \in \mathcal{A}_X \otimes \mathcal{A}_Y$, then $((x, y) \longmapsto \mathbb{1}_A(x) \mathbb{1}_B(y)) \in \mathcal{H}$.) $\square$

Theorem 10.10.2 (Fubini-Tonelli) Let $f : X \times Y \longrightarrow [0, +\infty]$ be an $\mathcal{A}_X \otimes \mathcal{A}_Y$-measurable mapping. Then the mapping

$$(x \in X) \longmapsto \int_Y f(x, y) \mu_Y(\mathrm{d}y)$$

is $\mathcal{A}_X$ measurable. Moreover, there exists a unique σ-finite measure $\mu_X \otimes \mu_Y$ on $X \times Y$ such that

$$\begin{aligned}\int_{X \times Y} f(x, y) (\mu_X \otimes \mu_Y)(\mathrm{d}(x, y)) &= \int_X \int_Y f(x, y) \mu_Y(\mathrm{d}y) \mu_X(\mathrm{d}x) \\ &= \int_Y \int_X f(x, y) \mu_X(\mathrm{d}x) \mu_Y(\mathrm{d}y).\end{aligned}$$

Proof
(1)Let

$$\mathcal{H} = \{^{\text{bounded } \mathcal{A}_X \otimes \mathcal{A}_Y\text{-measurable}}_{f: X \times Y \longrightarrow \mathbb{R}_{\geq 0}} \mid (x \in X) \longmapsto \int_Y f(x,y)\mu_Y(\mathrm{d}y) \text{ is } \mathcal{A}_X\text{-measurable}\}$$

Then $\mathcal{H}$ is a λ-family and $\mathcal{H}$ contains $\mathcal{C}$. So $\mathcal{H}$ contains all $\mathcal{A}_X \otimes \mathcal{A}_Y$ measurable bounded non-negative mappings.
(2) Let $\mathcal{A}_0 = \{A \times B \mid (A, B) \in \mathcal{A}_X \times \mathcal{A}_Y\}$. $\mathcal{A}_0$ is a semiring on $X \times Y$, $\mu_X(A) < +\infty$, $\mu_Y(B) < +\infty$. $\mu : \mathcal{A}_0 \longrightarrow [0, +\infty[$.

$$\nu(A \times B) \longmapsto \mu_X(A)\mu_Y(B) = \int_X \int_Y \mathbb{1}_{A\times B}(x,y)\mu_Y(\mathrm{d}y)\mu_X(\mathrm{d}x)$$

is a σ-additive mapping (by Beppo Levi). So ν extends in a unique way to a σ-finite measure on $\sigma(\mathcal{A}_0) = \mathcal{A}_X \otimes \mathcal{A}_Y$. For $f \in \mathcal{C}$,

$$\int_{X\times Y} f(x,y)\mathrm{d}\nu \int_X \int_Y f(x,y)\mu_Y(\mathrm{d}y)\mu_X(\mathrm{d}x).$$

So the same equality holds for any bounded non-negative $\mathcal{A}_X \otimes \mathcal{A}_Y$-measurable mappings $X \times Y \longrightarrow \mathbb{R}_{\geq 0}$. For general case, take an increasing limit. □

Theorem 10.10.3 (Fubini) Let $f : X \times Y \longrightarrow \mathbb{R}$ be an $\mathcal{A}_X \otimes \mathcal{A}_Y$-measurable mapping. f is $\mu_X \otimes \mu_Y$-integrable if and only if

$$\int_X \int_Y |f(x,y)|\mu_Y(\mathrm{d}y)\mu_X(\mathrm{d}x) < +\infty.$$

Moreover, when f is $\mu_X \otimes \mu_Y$ integrable, one has

$$\int_{X\times Y} f(x,y)\mathrm{d}(\mu_X \otimes \mu_Y) = \int_X \int_Y f(x,y)\mu_Y(\mathrm{d}y)\mu_X(\mathrm{d}x).$$

Proof
(1) f integrable $\Leftrightarrow |f|$ is integrable.
(2)

$$f = \sup\{f, 0\} - \sup\{-f, 0\}.$$

Apply Fubini Tonelli to $\sup\{f, 0\}$, $\sup\{-f, 0\}$. □

Definition 10.10.4 Let $(a, b) \in \mathbb{R}^2$, $a < b$, $\varphi : [a, b] \longrightarrow \mathbb{R}$. Assume that φ is the difference of two right continuous increasing mappings φ_1 and φ_2. For

$x \in \,]a, b]$, let

$$\varphi(x-) := \lim_{h>0, h\to 0} f(x-h).$$

Let $\Delta\varphi(x) = \varphi(x) - \varphi(x-)$. If $f : [a, b] \longrightarrow \mathbb{R}$ bounded Borel measurable, let

$$\int_a^b f(t)\mathrm{d}\varphi(t) := \int_a^b f(t)\mathrm{d}\varphi_1(t) - \int_a^b f(t)\mathrm{d}\varphi_2(t).$$

Theorem 10.10.5 Let $(a, b) \in \mathbb{R}^2$, $a < b$. Let φ and ψ be mappings from $[a, b]$ to $\mathbb{R}$, that can be written as difference of increasing right continuous mappings. Then

$$\int_a^b \varphi(t)\mathrm{d}\psi(t) + \int_a^b \psi(t)\mathrm{d}\varphi(t) = \varphi(b)\psi(b) - \varphi(a)\psi(a) + \sum_{t\in]a,b]} \Delta\varphi(t)\Delta\psi(t).$$

Proof We assume that φ and ψ are increasing.

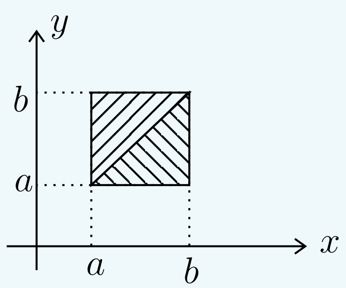
$$\int_a^b (\varphi(y) - \varphi(a))\,\mathrm{d}\psi(y) = \int_a^b \int_a^y \mathrm{d}\psi(x)\mathrm{d}\psi(y) = \int_{]a,b]^2} \mathbb{1}_{\{(x,y)\in]a,b]^2 | x\leq y\}} \mathrm{d}\varphi \otimes \mathrm{d}\psi.$$

Taking the sum, get

$$\begin{aligned} &\int_a^b \varphi(y)\mathrm{d}\psi(y) - \varphi(a)\left(\psi(b) - \psi(a)\right) + \int_a^b \psi(x)\mathrm{d}\varphi(x) - \psi(a)\left(\varphi(b) - \varphi(a)\right) \\ &= \int_{]a,b]^2} \mathrm{d}\varphi \otimes \mathrm{d}\psi + \int_{]a,b]^2} \mathbb{1}_D \mathrm{d}\varphi \otimes \mathrm{d}\psi, \end{aligned}$$

where $D = \{(x, x) \mid x \in \,]a, b]\}$. The first term is $(\varphi(b) - \varphi(a))(\psi(b) - \psi(a))$, and the second term is

$$\int_{]a,b]} \int_{]a,b]} \mathbb{1}_D(x, y)\mathrm{d}\varphi(x)\mathrm{d}\psi(y) = \sum_{x\in]a,b]} \Delta\varphi(x)\Delta\psi(x).$$



□